The psychological impact of dental enamel opacities

David Michael Williams BSc. (Hons)

Thesis submitted in candidature for the degree of
Doctor of Philosophy.
Cardiff University.

March 2005

Acknowledgements

First and foremost I would like to thank my principal supervisor, Dr Ivor Chestnutt for his constant support and advice. His patience and guidance made this thesis possible and allowed me to develop research skills that will help me long after the submission of this thesis. I would also like to warmly thank my second supervisor, Professor Paul Bennett. His good humour and suggestions have helped enormously over the last three years, and this thesis would look very different without his input.

There are a number of other people that I must also thank. Dr. Kerry Hood, for her brilliant statistical advice and guidance, and for answering a thousand questions. It was Kerry who suggested which statistical techniques would be most suitable for our often complex analysis. She also provided access to the necessary statistical software, and ran the tests on our data in the second and third experiments.

I would also like to thank the Psychology Department of the University of Wales, Swansea, who provided us with the laptop computer and other hardware and software we needed to create our second and third experiments. In particular Phil Heard spent a great deal of his time programming our design, and was always patient and helpful no matter how often I asked him to change it. Dr Rob Lowe was also always ready to give advice when asked.

Frank Hartles and the rest of the staff at the Dental Illustration Unit (both past and present) provided the photography equipment and expertise that allowed us to produce the digitally manipulated images that were vital to all three of the experiments in this thesis. The production of these images took longer than we anticipated, so thank you for not giving up on it.

I also have to say "thank you" to my generous sponsors, The Borrow Dental Milk Foundation. In addition to their funding they were also very supportive when I wished to take this research in a different direction to that initially intended.

Last, but certainly not least, I'd like thank my parents, Mary and Mike, and my girlfriend, Jane, for their emotional support, especially at those times when I could have torn my hair out with frustration!

My warmest thanks to all of you.

The psychological impact of dental enamel opacities

Summary

Introduction: Defects in the structure of tooth enamel may arise for a number of reasons, including excess exposure to fluoride during the period of tooth formation, so-called fluorosis. This results in altered tooth appearance, ranging from mild white spots to brown discolouration. The public's perception of fluorosis has not been adequately investigated. While previous research has considered the topic in terms of dental aesthetics, the psychological perspective has been largely ignored. Implicit measures of attitude have been used in psychological studies where the respondent may be unwilling (or unable) to answer explicitly, but have not been applied to a study of dental appearance.

Method: This thesis describes three experiments which sought to identify personal characteristics attributed to those with various dental conditions. Standardised intra-oral or extra-oral images were used to avoid the effect of factors such as tooth size and shape, and variation in facial form on judgements made by the participants. An implicit attitude measure, based on the Affective Priming Task, was developed and utilised in two of the experiments to allow both self-reported (explicit) and implicitly measured attitudes to be assessed.

Results: The results showed that mild and moderate fluorosis were not judged differently to normal enamel when viewed in an extra-oral image, but severe fluorosis and dental caries were judged less favourably. This was true whether attitudes were measured explicitly or implicitly.

Cueing participants by asking them to look at the mouth increased the negative impact of severe fluorosis. Asking participants to attribute characteristics based on intra-oral images further increased the impact of severe fluorosis and also resulted in mild fluorosis being judged less favourably than normal enamel.

Conclusion: Implicit measures of attitude can be used to identify variation in the strength of attributions made about different dental conditions. Explicit measures may measure the valence of attitudes. This study suggests that mild fluorosis may be perceived more favourably than untreated dental caries. It has also shown that traditional approaches to assessing the perception of fluorosis may have exaggerated its impact.

Table of contents

List of Tables

List of Figures

Chapter 1. Literature Review

1.1. Introduction

This literature review is divided into three distinct sections. The first is concerned with the importance of physical appearance and the consequences of deviating from a normal appearance. This section concludes with a review of how dental factors affect appearance. The second section of the literature review chronicles the identification of dental fluorosis, its aetiology, prevalence, and how it is perceived. Finally the literature on the measurement of attitudes, and which methods of attitude measurement might be applicable to the public's perception of dental fluorosis is reviewed. Implicit measures of attitude in particular are considered from the perspective of the current thesis. The details of how the literature search was conducted are given in Appendix A.

1.2. Appearance and disfigurement

1.2.1. The social importance of appearance

The way we look has been important in human social interaction since ancient times and there is evidence of the importance of physical appearance from a rich variety of sources (Rumsey, 1997). Previous reviews have concluded that in the past there has been a reluctance to acknowledge the true impact of physical attractiveness (Berscheid, 1980) and that social attributions are made based on appearance, which in turn effect behaviour and personality development. Helping people to improve their appearance can promote positive socialization experiences and healthy personality development (Adams, 1980).

It is clear that being physically attractive can be an advantage in many ways. In fact Dion, Berscheid, & Walster (1972) found a "what is beautiful is good" stereotype, where attractive people were assumed to have more socially desirable traits than unattractive people and to lead better lives. However, a subsequent meta-analysis (Eagly, Makhijani, Ashmore, & Longo, 1991) revealed that the "what is beautiful is good" stereotype is not quite as pervasive as this phrase suggests, since perceptions of integrity and concern for others were not linked to physical attractiveness. Nevertheless, the meta-analysis did find that perceptions of potency, adjustment, intellectual competence, and, in particular, social competence were all associated with physical attractiveness.

In another meta-analysis, Langlois, Kalakanis, Rubenstein, Larson, Hallam, & Smoot (2000) found that "both within and across cultures, people agreed about who is and is not attractive... (and that) attractiveness is an advantage in a variety of important, real-life situations" (p.399). Their findings also suggest that attractiveness is as important for males as it is for females, and for children as it is for adults. Furthermore, they did not find that familiarity had any effect on attractiveness. In other words, the effects of attractiveness are as strong for people you know well as they are for strangers. These findings contradict certain maxims that they were considering during their analyses, these being "beauty is in the eye of the beholder", "never judge a book by its cover", and "beauty is only skin deep". Apparently these maxims about attractiveness understate its true importance.

Many of the studies described below highlight how being physically attractive has been shown to be an advantage in terms of how others perceive and treat us, and also describe some of the links established between physical attractiveness and certain favourable personality traits.

However, before considering the wealth of evidence on the importance of appearance it should be noted that there is some counter evidence to the idea that looks are all important.

Feingold (1992a) conducted a meta-analysis on mate selection research in order to investigate evolutionary theories derived from Triver's parental investment model (Trivers, 1972; Trivers, 1985). This model contends that women are more likely than men to seek a mate who possesses characteristics unrelated to physical attractiveness, and which serve as cues to resource acquisition in order to maximise the survival or reproductive prospects of their offspring. As predicted, women accorded more weight than men to socioeconomic status, ambitiousness, character (i.e. honesty, sincerity), and intelligence, with the largest gender differences being observed for cues to resource acquisition (status and ambitiousness). Gender differences were not expected or found in preferences for characteristics unrelated to progeny survival (in this case sense of humour and "personality", where personality refers to charisma and expressiveness). Similar results were found when these findings were compared cross-culturally, cross-generationally, and using different research paradigms, although the research available for these comparisons was limited. Furthermore, Livingston (2001) developed the Perceptual Reliance Index (PRI) which is a measure of the extent to which an individual makes social judgements based on physical cues. His findings showed that while some individuals do make different social evaluations of attractive and unattractive people, others do not.

Returning to the evidence on the importance of appearance, Rumsey (1998) noted there were many studies conducted in the 1960's and 1970's that found advantages in being attractive (although relatively few papers have been published on this topic since that time). These papers tended to identify attractiveness as being beneficial without developing a clear theoretical framework beyond support for Dion et al.'s (1972) "what is beautiful is good" stereotype. Examples of the evidence from the1960's and 1970's include

Miller (1970), who found attractiveness was associated with positive traits and unattractiveness associated with negative traits, when people were asked to make judgements based on photographs of others. In addition to this general association between attractiveness and the attribution of traits, appearance has been shown to impact on a variety of situations important to people's lives, including education, court-room appearances, wage level, interpersonal attraction and even long term relationships. These studies are now described.

The importance of appearance is not just relevant to adults. Clifford & Walster (1973) found that how attractive children were affected teacher's expectations of how intelligent the child was, how interested in education the child's parents were, how far the child was likely to progress in school, and how popular the child would be with his/her peers. The more attractive children received more favourable expectations than the unattractive children. However, in a similar study Shaw & Humphreys (1982) found that neither overall facial attractiveness nor dental appearance effected teachers' ratings of pupils. While both of these studies asked teachers to base their assessment on photographs and report cards of the pupils, other studies have looked at the effect of appearance on teachers ratings of children they have actually taught. Dare (1992), asked teachers to rate the appearance, and several other qualities of children they had taught for 3 months. She reported that the children who were rated as more attractive were also rated more favourably for a number of other qualities including intelligence, behaviour, popularity, and as having parents who are more interested in education. A study by Salvia, Algozzine, & Sheare (1977) found teachers gave better grades to attractive children, even after teaching them for a full academic year, although in an objectively assessed test, attractive children only scored better than unattractive children in one of the three grades considered. The effect of children's appearance on teacher ratings may be cross-cultural since the Dare (1992) study was conducted in Nigeria, while the Salvia et al. (1977) study was run in America.

Children are not only judged on their appearance by their teachers, their peers also make such judgements. Even children between the ages of three and six-and a half years old showed a preference for attractive adults and children, and inferred that attractive children were more likely to behave prosocially, while unattractive children were perceived as more likely to be antisocial (Dion, 1973).

Another study of appearance in the 1970's showed that in a mock trial, attractive defendants were less likely to be found guilty, and juries recommended they receive less severe punishments than unattractive defendants (Efran, 1974). Physical appearance has also been linked to the wage level and wage growth of male and female workers (Loh, 1993), although this study considered only the height and weight of the workers, rather than measurements of physical attractiveness.

Appearance may also dictate the extent to which other people notice us. Maner, Kenrick, Becker, Delton, Hofer, Wilbur, & Neuberg (2003) found that when viewing images of faces under conditions which limited attentional capacity, both male and female participants attended to, and recognised previously seen faces to a greater extent when viewing attractive female faces than they did when viewing unattractive or average-looking female faces (the degree to which faces were attended to was measured by participants' estimates of the frequency of attractive faces and by tracking eye movements). When viewing male faces there were different trends for men and women. Male participants did not pay more attention to, or more easily recognise attractive male faces. Female participants did pay more attention to attractive male faces but did not recognise them more easily than they recognised less attractive male faces. This suggests being physically attractive is beneficial for both men and women when looking for a partner, but is especially beneficial for women.

A study by Berry & Miller (2001) demonstrated the importance of appearance in inter-personal attraction by investigating interactions between opposite sex strangers. They found that such interactions are affected by the physical attractiveness of females but how attractive males are is less important. Berry & Miller (2001) surreptitiously filmed 51 unacquainted male-female couples for six minutes, and then asked the participants and neutral observers to rate the interaction. Additionally, participants' personality traits were measured on the Big Five personality model. This model is based on the idea that personality can be described and measured on five broad dimensions: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism (McCrae & Costa Jr., 1997). They found women's physical attractiveness –but not their personality scores- predicted their own, the male stranger, and independent observers' evaluations of the interaction. Conversely, men's personality scores –extraversion in particular- predicted their own and observers' ratings of the interaction. This suggests that the womens' appearance was important to how satisfied men were with the interaction, but the mens' appearance was not important to how satisfied women were with the interaction, which is further evidence for models of attraction, such as the parental investment model (Trivers, 1972; Trivers, 1985, described above), that propose women value physical attractiveness in their partner less highly than men do.

Physical attractiveness may also influence satisfaction with long-term relationships, such that those with more attractive partners are more satisfied (Sangrador & Yela, 2000). However, in this study the physical attractiveness of the partner was only rated by the respondent, it is therefore possible that they rated their partners as being more physically attractive because they were satisfied with them. Without an objective measure of attractiveness it is unclear what impact physical appearance actually had on satisfaction with the relationship. Moreover, while physical attractiveness was found to be the most important factor in short-term, sporadic relationships, it was not as important to those in long-term, committed relationships, although it could influence the manner in which people fell in love and was linked to feelings and thoughts associated with love.

Long-term relationships may also influence attractiveness ratings of people outside the relationship. Simpson, Gangestad, & Lerma (1990) found that people in relationships rate members of the opposite sex as being less attractive than do single people, but they do not rate same-sex others, or opposite-sex old people differently to single people. This suggests that in order to protect their relationship people may, either deliberately or unintentionally, play down the attractiveness of possible alternative partners. So, while physical attractiveness may influence social interactions in many ways, its impact might be reduced if the perceiver is in a stable relationship.

1.2.2. What is considered physically attractive?

Given that it has already been noted that what is physically attractive is consistent both within and across cultures (Langlois et al., 2000), it is worth clarifying what is meant by physically attractive.

A series of studies which attempted to evaluate facial attractiveness by determining which facial feature correlated most strongly with overall facial attractiveness have provided mixed results. In a study of how female spectacle wearers self-ratings of attractiveness differed from contact lens wearers and participants with uncorrected vision, participants were asked to rate their overall facial attractiveness and then to rate the attractiveness of the individual features of their face. The rating of the mouth correlated most strongly with the overall rating of the face, followed by the eyes, hair, nose and facial structure -shape and complexion (Terry & Brady, 1976). A subsequent study found that when judges rated the attractiveness of photographs of other people when their faces were whole, and then rated their facial features individually (after the photograph was dissected into the component facial features) the mouth correlated most strongly with the rating of the whole face, followed by eyes, structure of face, hair and nose (Terry & Davis, 1976). However in a later study, Terry (1977) asked participants to rate the overall attractiveness of a photograph of a face and then to rate the component features when they were still in the context of the face rather than when they were isolated. This time eyes correlated most closely with the overall rating followed by nose, mouth, eyebrows, complexion, chin, and expression.

Since these inconclusive early studies, investigations of facial attractiveness have employed more sophisticated methodologies and become more theory driven, often using computer technology and tending to look at hypotheses that suggest attractiveness is determined by an appearance indicating health, which is typically thought to be denoted by symmetry (i.e. symmetrical facial features) and averageness (an appearance that was in no way unusual). These hypotheses, which include the "good genes" hypothesis, have been described by Zebrowitz, Fellous, Mignault, & Andreoletti (2003) and often consist of several similar theories that have been grouped together. It is beyond the scope of this thesis to review each of these hypotheses and theories individually, but many of them share the underlying theme that certain traits, such as symmetry and averageness, are found attractive because it they were adaptive evolutionary preferences, or because it was adaptive to avoid partners who displayed the opposite of these traits (i.e. asymmetrical features or a nonaverage appearance). Therefore, it is sufficient, for the purpose of this review, to discuss studies that attempted to identify what physical characteristics are considered attractive without going into great depth as to why they are considered attractive (although theories will be mentioned where relevant).

In one such study, Grammer & Thornhill (1994) generated computer images of mens' and womens' faces, and of composites of the faces of each sex, and asked people to rate opposite sex faces. As they predicted, men preferred symmetry and averageness in women's faces although when symmetry was partialed out of the ratings, averageness did not correlate with the men's ratings of women. Women preferred symmetry in men's faces, as predicted, but averageness had a negative effect on women's ratings of men which was unexpected. Men rated prominent cheekbones as sexy in women. However, they also rated small eyes as attractive which is the opposite of what was expected, and which, as the authors note, showed:

> Inconsistency with... previous studies... It is unclear... why we obtained this result, but it could reflect methodological variation between our study and other studies or a finding peculiar to our sample (p.240).

Women found a large jaw and a wide mouth attractive in men and also preferred a broader face. Prominent cheekbones were positively but not significantly correlated with attractiveness in women's ratings of men.

A study of symmetry and averageness in appearance conducted by Rhodes, Zebrowitz, Clark, Kalick, Hightower, & McKay (2001) investigated whether facial symmetry and an average facial appearance are perceived as being healthier than asymmetrical or distinctive faces. They based this on suggestions that features which indicate good health, and therefore good reproductive potential, are attractive to the opposite sex (they cite meta analyses by Langlois et al., 2000, and Feingold, 1992b, as showing a weak link between health and attractiveness). They found that both symmetry and averageness are perceived as being healthier and that averageness does indicate good health to some degree. However, no such relationship was found between symmetry and health, which contradicts the "good genes" hypothesis for why symmetry is regarded as attractive.

Continuing with the theme that attractiveness is determined by features which denote health, Zebrowitz et al. (2003) suggested that the less your appearance suggests you have a genetic anomaly the more attractive you are to the opposite sex. They found impressions of the traits of normal adults were predicted by the similarity of their appearance to people with genetic anomalies. The similarity of their appearance to an individual with a genetic anomaly was assessed by the extent to which a neural network trained to recognise anomalous faces confused their faces with anomalous ones. In order to create the neural network Zebrowitz et al. took photographs of a number of faces of individuals with genetic anomalies and marked key points on each face based on the positioning of the facial features (e.g. the tip of the nose). Using information about the distance between these points a neural network was created to recognise the faces of adults with birth defects and deformities. A second neural network designed to recognise babies was created using the same methodology. The extent to which the neural network mistook normal adult faces for an individual with a genetic anomaly predicted human judges' impressions of attractiveness, sociability, warmth, health strength, and intelligence. The extent to which the neural network mistook normal faces for anomalous faces did not predict the actual traits of the normal faces for health or intelligence (not all of the trait measurements were available for all of the people who supplied the normal faces). These findings supported the anomalous face overgeneralization hypothesis i.e. that faces that look anomalous are assumed to share the traits of anomalous individuals, and that therefore humans prefer faces which apparently have "good genes" as they signify a better chance of being able to reproduce. The fact that the neural network did not predict actual health or intelligence only perceived health and intelligence does not support the "good genes" hypothesis because the individuals considered attractive did not appear to actually possess better genes than unattractive individuals

The neural network trained to recognise babies' faces predicted impressions of babyfaceness in normal adults. This supported the baby face overgeneralization hypothesis i.e. that faces that look like babies' faces are assumed to share the traits of babies; that is, they are assumed to be warm, physically weak, submissive and naive. The neural networks identified elderly adults as being more anomalous and more babyfaced than young adults although human judges rated the elderly as less babyfaced. The neural networks identified overweight faces as being more babyfaced but not as being more anomalous, but human judges' rated overweight people as more anomalous but not more babyfaced. This suggests that some attributions that are made are based on cultural rather than genetic factors.

1.2.3. Appearance and personality

Being attractive may have a double influence on personality. Firstly there is evidence that people make attributions about the personality of others based on their appearance, and that they make largely positive attributions about attractive people (Section 1.2.1); and in addition to this being physically attractive may also lead to certain social advantages from within our own personality. For example attractiveness has been linked to certain positive personality traits such as assertiveness. In a study by Jackson & Huston (1975) attractive females were shown to be more assertive than unattractive females when responding to impolite behaviour (they interrupted the experimenter after a shorter period of time than unattractive participants when kept waiting), although there was no difference between the attractive and unattractive females according to their self-ratings of assertiveness. Furthermore, attractive college students were shown to be more likely to resist peer pressure influences and to have internalized socially desirable personality characteristics (Adams, 1977).

However, while physically attractive people may be attributed with many desirable personality traits they do not necessarily possess all of them. Feingold (1992b) conducted a meta-analysis on the relationship between the attractiveness stereotype and actual attractiveness. He found that attractive people were perceived as more sociable, dominant, sexually warm, mentally healthy, intelligent, and socially skilled than physically unattractive people. However, the correlational literature indicated generally trivial relationships between physical attractiveness and measures of personality and mental ability, although attractive people were less lonely, less socially anxious, more popular, more socially skilled, and more sexually experienced than unattractive people.

Interestingly, Feingold (1992b) found a relatively low correlation (.24) between physical attractiveness attributed by others and self-rated physical attractiveness. Personality dimensions such as dominance, emotional stability, and self-esteem were positively correlated with self-rated attractiveness but were effectively unrelated to physical attractiveness attributed by others. In contrast, social skills and freedom from public self-consciousness were correlated with other-rated attractiveness but not with self-rated attractiveness. Measures related to social interactions (i.e. freedom from social anxiety and loneliness, popularity with the opposite sex, sexual experience) were correlated with both measures of attractiveness and academic ability was associated with neither. It is therefore possible that some of the advantages of being attractive are available to everybody, as long as they believe that they are attractive.

Feingold's (1992b) finding that relationships between physical attractiveness and measures of personality and mental ability tend to be trivial suggest that that the personality traits which appear more common in physically attractive people are likely to have been developed rather than being innate to physically attractive people. That is, attractive people appear to be more assertive, etc. because they have experienced more positive social interactions than less attractive people, and not because they are naturally possess a different type of personality, if a different type of personality were the explanation then a stronger relationship would be expected between physical attractiveness and personality. However, recent work by Zebrowitz and her co-workers suggests that while personality traits may not be innate to physically attractive individuals *per se*, there may be a complex relationship between personality and appearance over the lifespan.

In a series of experiments Zebrowitz and her colleagues utilised longitudinal, archival data, including personality tests and photographs of subjects taken at a number of different ages, to investigate the relationship between personality and appearance across the life-span. They asked participants to rate the extent to which the photographed individuals appeared to possess various characteristics and then looked at how the perceived ratings correlated with the archival measures of that characteristic. Their work, which is described below, found support for a number of theories including the self-fulfilling prophecy, self-defeating prophecy, the Dorian Gray effect and the artifice effect. For example, Zebrowitz, Collins, & Dutta (1998) investigated how stereotypes associated with different types of appearance interact with personality, and reported that the relationship between personality and appearance was "complex, varying as a function of age, gender, and the appearance quality under investigation" (p.745).

They developed a measure of the prototypically attractive personality (the PAP scale) and the prototypically babyfaced personality (the PBP scale).

This was done by asking expert judges to sort Q-sort items into categories based on the responses they believed a stereotypically physically attractive (or babyfaced) person would make (in a typical Q-sort participants put statements and traits into categories based on how characteristic the items are of themselves). They identified the prototypically attractive personality as being sociable, dominant, non-hostile, and dependable, and the prototypically babyfaced personality as being sociable, submissive, non-hostile, and non-dependable. Then they correlated archival Q-sort data from subjects with the data from the prototypical personality types to create an index of how similar the subjects' personality was to that of a prototypically attractive or babyfaced individual. They found that stereotypes associated with different types of appearance could lead to self-fulfilling prophecies (e.g. high attractiveness in adolescence and the thirties was predictive of a more prototypically attractive personality for men in their fifties), self-defeating prophecies (e.g. higher babyfaced appearance in boys during childhood and puberty predicted a less prototypically babyfaced personality in adolescence), and a Dorian Gray effect (whereby appearance grows to reflect personality, e.g. women who had a prototypically attractive personality in adolescence and their thirties, were more attractive in their fifties) thereby demonstrating that appearance and personality can interact and influence each other.

Further evidence of the influence of early personality on appearance comes from Zebrowitz, Voinescu, & Collins (1996) who reported an artifice effect (i.e. appearance develops in such a way to disguise true personality) in honesty and perceived honesty for women (but not men), such that more dishonesty early in life predicted a more honest appearance later in life for women. As in the Zebrowitz et al. (1998) study, this was achieved by using archival data (in this case photographs and scores from an Honesty Q-sort at five age levels: childhood, puberty, adolescence, 30-40 years old, and 50-60 years old) and then asking judges to rate the photographs for how honest and attractive they appeared. Their findings imply it is possible, in some cases, to adapt appearance to achieve certain aims, such as successful lying. They also found that physical attractiveness impacted on how honest people were perceived to be, such that attractiveness (as well as babyfaceness, facial symmetry and large eyes) had positive, independent effects on perceived honesty, which is further evidence of the global benefits of an attractive appearance.

It has been shown that the attributions people make about the personality of others based on their appearance can sometimes be accurate. Zebrowitz, Hall, Murphy, & Rhodes (2002) found that people can judge the intelligence of strangers from facial photographs with above chance accuracy (although there was only a modest effect). Furthermore, facial attractiveness contributed to perceivers' accuracy. In other words being attractive may lead people to attribute you with intelligence and may even be linked to intelligence. However, this is a somewhat tentative finding given the modest effect sizes and the fact that at some age levels (adolescence or later adulthood) there was no significant relationship between perceived intelligence and measured intelligence. Furthermore, due to missing photographs and IQ scores the samples were not identical across the lifespan. In addition to their own study Zebrowitz et al. (2002) conducted a meta-analysis on previous studies of perceived and measured intelligence that used facial photographs and IQ score (they noted other studies have used different measures of intelligence). The meta-analysis found a significant positive correlation between perceived and measured intelligence, although it should be noted the studies in the meta-analysis were very old, many of them from the 1920's and 1930's. Their findings also contradict the meta-analysis of Feingold (1992b) who found no relationship between self- or other-rated attractiveness and academic ability. Overall it appears there is evidence of a link between perceived intelligence and measured intelligence, but that more research is needed to confirm this relationship.

Zebrowitz, Olson, & Hoffman (1993) investigated how stable physical attractiveness and babyfaceness were across the lifespan; this is of interest since stability of appearance has implications for the effect of appearance on personality. They found that attractiveness had differential stability across the lifespan, i.e. how attractive you are considered among your peers remains constant over your lifespan. Babyfaceness showed differential stability for males in childhood and throughout their 30's, and for females throughout adolescence. This suggests that the stereotypes associated with attractiveness will remain relevant (or not relevant for unattractive individuals) to an individual when dealing with members of the same peer group throughout the course of his or her life, but that stereotypes associated with babyfaceness will only be pertinent during certain periods of his or her life.

The novel nature of Zebrowitz and colleagues work means it needs to be replicated before its full impact can be assessed, but it does appear to have shown some interesting results regarding the relationship between appearance and personality.

1.2.4. The consequences of disfigurement

As Rumsey (1998) noted, studies on "the effects of physical attractiveness abounded in the 1960's and 1970's... (but) the effects of ugliness and disfigurement have been much less fashionable (as a research topic)" (p.576-7). However it is worth considering the work that has been done in this area to highlight the duality of the effects of appearance. The benefits of a pleasing appearance have already been noted (Section 1.2.1), however, an appearance that deviates from the norm has more severe consequences than merely the absence of these benefits. Disfigurement, even mild disfigurement, can result in negative reactions from other people and problems with self-image.

This has been demonstrated by Demelleweek, Humphris, Hare, & Brown (1997) who showed children video clips of other children, some of who were made up to appear to have facial port-wine stains. The participants were then asked questions about the children in the video clips. The participants expected that port-wine stains would attract staring and teasing. They thought it would cause self-pity for the boy with the stain, and that it would make getting a boyfriend more difficult for one of the two girls with a port-wine stain. However, the stain did not significantly affect the participants' assessment of the attractiveness and character or willingness to interact with/befriend the children, although "it is possible that the participants were somewhat reluctant to express their own prejudice but recognized it would occur" (p.483).

Appearance may also affect how an individual perceives themselves. For example, Lister (2001) noted that disfigured patients may have difficulty with their feelings of confidence and self-worth. Furthermore, in their review of the literature concerning the current psychological understanding of adjustment to acquired and congenital disfiguring conditions, Thompson & Kent (2001) discuss the feelings of disfigured patients, which include anxiety, shame and depression.

These feelings are not limited to severely disfigured people. Moss (1997) and Carr, Harris, & James (2000) both noted that the distress felt by the disfigured individual could be disproportionate to the degree of apparent disfigurement, with even relatively minor blemishes causing great suffering. It may even be the case that people with minor disfigurements can suffer greater psychological distress than the more obviously disfigured. Lansdown, Lloyd, & Hunter (1991) assessed 27 disfigured children (grouped into 3 categories by severity of disfigurement, measured by parental reports of how often they were stared at), and 12 controls, plus 26 siblings of the experimental group and 12 control siblings. They found mildly disfigured children had lower self-esteem than more severely disfigured children, although the effect did not reach significance. There were also no significant differences between the disfigured group and the control group for self esteem. The parents of the children also viewed mildly disfigured children as having more disturbed behaviour, although again, any between group differences did not reach statistical significance.

Possible explanations for the distress that can be felt by those who have a minor disfigurement are that it is more socially acceptable to make them the subject of jokes and tease them than it is to mock people with severe disfigurements. The mildly disfigured may also experience greater uncertainty about what type of reaction their disfigurement will provoke, or whether it will be noticed at all, resulting in higher levels of social stress and discomfort. In contrast, those with a serious disfigurement know that their appearance will be noticed and can prepare themselves accordingly, however, that isn't to suggest that severely disfigured people do not suffer as a result of their appearance, (Macgregor, 1970; Macgregor, 1990).

1.2.5. The social importance of dental appearance

In a review of the literature concerning the social and psychological implications of dentofacial disfigurement Macgregor, (1970) stated:

> "there are two... handicapping aspects associated with dentofacial deformity... the area around the mouth is both emotionally charged and strongly connected with one's self image... (and secondly) such defects interfere with the flow of social interaction" (p.233).

Furthermore, early research indicated that dental appearance was important to the public (Linn, 1966) and that, as noted previously, the mouth was possibly the most important individual facial feature in terms of assessing attractiveness (Terry & Brady, 1976; Terry & Davis, 1976). Subsequent research found that a good dental appearance is regarded by the public as a requirement for some prestigious occupations (Jenny & Proshek, 1986), although in a recent study Richmond, Tatarunaite, Playle, Hood, & Shaw (in press) found that no single facial feature was especially important in overall attractiveness and that teeth in particular had little impact on attractiveness. More evidence that a dental appearance which differs from the norm has potential social consequences was provided by Roberts-Harry, Hathorn, & Stephens (1992). They found that children with cleft palate or lip were still considered less attractive than children without cleft palate or lip, even after the cleft had been repaired.

Dental appearance has a considerable impact on how others perceive us. Eli, Bar-Tal, & Kostovetzki (2001) found that "individuals with normal dentition were consistently scored as being more aesthetic and more successful socially and professionally than individuals whose appearance was altered to show decayed teeth" (p.133). Furthermore, the effect of dental appearance was more prominent on the physical attractiveness stereotype when participants evaluated members of the opposite sex than when they rated members of their own sex.

It has been found, by showing different dentofacial arrangements within otherwise standardised photographs of individuals, that children with a normal dental appearance are judged to be better looking, more desirable as friends, more intelligent, and less likely to behave aggressively than children with unusual dental appearance (Shaw, 1981). Using similar methodology it has also been shown that normal dentofacial appearance in young adults contributes positively to judgements made about a variety of personal characteristics although overall facial attractiveness is more influential than dental appearance (Shaw, Rees, Dawe, & Charles, 1985).

Dental appearance can also affect judgements made about some personal characteristics. Feng, Newton, & Robinson (2001) found that decayed or discoloured teeth were associated with perceived lower levels of adjustment and intellectual competence. In a similar study, Newton, Prabhu, & Robinson (2003), found that dental decay had a significant negative impact on participants' appraisals of the social competence, intellectual achievement, and psychological adjustment of others. Like Feng et al. (2001) they asked participants to rate standardised images in which the dental appearance was digitally manipulated to simulate decayed teeth. Furthermore, they suggested future research explore the effects of other dental conditions on participants' perceptions, which suggests that they believe decay is not the only condition which can impact on observers' judgements.

Disfigured dental appearance can be a cause of distress. One study found that 7% of children were teased about their dental appearance, that comments about the teeth were more likely to cause upset than teasing about other features, and that children who were teased about their teeth were twice as likely to suffer harassment than children who were teased about other things. Furthermore, the more deviant the dental appearance, the more salient it became (Shaw, Meek, & Jones, 1980). It has also been shown that elderly patients reported improved self-image and social interaction as a result of dental treatment (Fiske, Gelbier, & Watson, 1990), which suggests the dental appearance of these patients was causing them some upset before it was corrected. There is clearly a wealth of evidence that dental appearance can impact on overall appearance, and by doing so impact on peoples' wellbeing. Furthermore, it has been shown that standards of dental appearance may be similar across cultures. Cons, Jenny, Kohout, Freer, & Eismann (1983) found that participants in Australia, the former German Democratic Republic, and the U.S.A. showed similar preferences in dental appearance.

1.2.6. Appearance and the current research

The present study considers the social and psychological impact of dental fluorosis, a condition that causes discolouration of the teeth. In order to emphasise the relevance of this study, the literature on appearance and disfigurement, and specifically the importance of dental appearance, has been reviewed. Previous research has demonstrated that physical appearance plays an important role in social interactions and self-image, and that an unpleasant appearance or disfigurement, even a minor disfigurement, can have a negative impact in both of these areas. Moreover, dental appearance is an important part of facial appearance. This literature justifies the current investigation into the social impact of dental fluorosis.

1.3. Fluorosis

1.3.1. Introduction

The story of fluorosis, its early description, identification of aetiology, and implications for oral health has been described in detail by Murray, (2003). At the beginning of the twentieth century, Frederick McKay, a dentist practising in Colorado Springs, Colorado, U.S.A. noticed that many of his patients' had malformed dental enamel, known locally as "Colorado stain". McKay spent over 30 years investigating this "mottled enamel", his research leading him to the conclusion it was only occurring in certain geographic areas, and that it was caused by an unknown agent in the water supply. The work of McKay came to the attention of H. V. Churchill, the chief chemist of ALCOA (a company specialising in aluminium products). Fearing that aluminium could be blamed for the mottling, he arranged for water from the sites listed by McKay to be tested for rare elements, and unexpectedly identified the presence of fluorine in the water, which was responsible what became known as dental fluorosis.

McKay's work paved the way for H. Trendly Dean to discover that fluoridated water not only caused enamel mottling, but also reduced the prevalence of dental caries. Furthermore, Dean deduced that the optimal level of fluoride in the water was one part per million (ppm). At this level the teeth had the maximum protection against caries, whilst the level of enamel mottling was minimised (Murray, 2003).

Dean's work in the 1930's, led to the proposal that fluoride be added to the public water supply to provide protection against tooth decay. A landmark study was begun in 1945, involving the fluoridation of the drinking water of Grand Rapids. This showed that 6 and-a-half years after the fluoridation commenced, the children of Grand Rapids had almost 50% less caries than the children of the non-fluoridated control town of Muskegon (Murray, 2003).

Since this study water-fluoridation has taken place in numerous countries throughout the world and has been shown to be effective in preventing dental caries. Throughout this time, fluorosis, as a consequence of adding fluoride to the water supply, has been a consideration. Although 1 ppm is recognised as the optimal level lower levels may be appropriate in tropical climates (Lesan, 1987; Murray, 2003).

Recent work has shown that even fluoridation at 1 part per million (ppm) can cause fluorosis that is of aesthetic concern (McDonagh, Whiting, Bradley, Cooper, Sutton, Chestnutt, Misso, Wilson, Treasure, & Kleijnen, 2000). Such public health concerns over fluorosis have led to a recommendation that the upper limit on the level of fluoride in the water supply in the Republic of Ireland be reduced to 0.8 ppm (Department of Health and Children, 1999). Therefore, public reaction to, or perception of, fluorosis is of interest to the dental profession and is also a wider public health concern as fluorosis may be considered a side effect of fluoridation and increased exposure to fluoride from this and other sources.

1.3.2. The aetiology of fluorosis

Fluorosis is caused by the effect of excess levels of fluoride on ameloblasts (the cells that produce enamel) during tooth formation. High levels of systemic fluoride result in defective enamel development and mineralization (Yaeger, 1976). This results in altered tooth appearance, which can range from mild white spots (hypomineralisation), which are barely noticeable, to unsightly brown discoloration and pitting (hypoplasia) (Murray, Rugg-Gunn, & Jenkins, 1976).

While it is apparent that fluorosis results from fluoride ingested during tooth formation, the precise point of tooth formation at which fluoride intake is critical in the development of fluorosis is unclear. Barsden (1999) identified when fluoride intake is most likely to cause fluorosis. He found that:

1) Children who were introduced to fluoride during the first two years of life were significantly more likely to have fluorosis than those introduced to it after they were two years old.
2) Children who had their fluoride exposure reduced after they were two years old were significantly more likely to have fluorosis than those that had it reduced before they were two years old.
3) Children who were exposed to fluoride for long periods (more than two out of their first four years) were significantly more likely to have fluorosis than children that were exposed to fluoride for short periods (less than two out of their first four years).

4) Neither the secretory phase early nor the maturation phase of enamel formation can be singled out as being the high risk period for developing fluorosis in the maxillary permanent central incisors.

Levy, Hillis, Warren, Broffitt, Mahbubul Islam, Wefel, & Kanellis (2002) found that 6-9 months-old was the most important period in fluorosis aetiology for the primary dentition. However, in the case of the permanent maxillary central incisors, where fluorosis would be most easily noticed, ingesting excess fluoride up to the age of approximately 30 months will make the teeth vulnerable to fluorosis. Evans & Stamm (1991) found the found the maxillary central incisors to be most at risk to fluorosis from dietary fluoride in a critical four-month period commencing around 22 months after birth, and in a later study Evans & Darvell (1995) found the maxillary central incisor to be most at risk of fluorosis from dietary fluoride between 15-24 months old for males and 21-30 months old for females. Despite, these data, while it is not yet certain which period of tooth formation is critical in the development of fluorosis, or even if there is a single period that determines if fluorosis will occur, the literature does suggest that the first two and-a-half years of life are crucial in the aetiology of fluorosis.

In addition to tooth formation, there are two further issues that have recently become noteworthy when considering the aetiology of fluorosis. These are sources of fluoride and genetic factors. Although there is little literature available on these areas it is worth outlining them as they may guide future research.

Although fluorosis is often debated in the context of water fluoridation schemes, Levy & Guha-Chowdhury (1999) noted that the total fluoride intake from all sources (not just drinking water) is important when investigating the aetiology of fluorosis. Among the other sources of fluoride they mention dentifrice, mouthrinse, fluoridated gels, and other fluoride supplements (tablets, drops, lozenges) as well as milk/infant formula, and beverages (tea, coffee, juice, carbonated drinks). Work by Riordan (2002) suggests that by reducing the use of these supplements it may be possible to reduce the prevalence of fluorosis.

Numerous reviews have also reached the conclusion that the misuse of dentifrice and fluoride supplements can result in fluorosis. Bowen (2002) felt the increasing prevalence of fluorosis was attributable to the "inappropriate use of fluoride containing products" (p.1407) such as supplements (in the form of lozenges, tablets or drops), toothpaste, and infant formula. Newbrun & Horowitz (1999) claimed "unintentional swallowing of fluoride-containing dentifrices... is a major factor in the increasing prevalence of enamel fluorosis" (p.533). This was supported by Levy, Kiritsy, & Warren (1995), who recommended that the fluoride content of foods, beverages and infant formula be monitored closely to limit excessive fluoride intake; that ingestion of dentifrice by young children should be controlled and it should be emphasised that they use only small quantities of dentifrice; furthermore, dietary fluoride supplements should be considered only for those children at higher risk for dental caries.

Horowitz (1992) recommended that toothpastes containing a lower level of fluoride (400-500 ppm rather than the typical 1000 ppm) be marketed for children. He believed that although the ingestion of fluoride toothpastes by pre-school children was not the major contributor to fluorosis, it was a risk factor, supporting the view of Ripa (1991) who believed that dentifrice was unlikely to be the cause of much fluorosis. However, Tavener, Davies, Davies, & Ellwood (2004) suggested that the severity of fluorosis may increase when using toothpastes with a higher concentration of fluoride; they found that the prevalence of fluorosis at the TF 3 level was higher for participants using a toothpaste with a fluoride level of 1450 ppm compared those using a toothpaste with a fluoride level of 440 ppm, although the overall prevalence of fluorosis was similar between the two groups.

While ingesting an excess of fluoride during the period of tooth development is widely accepted as the cause of fluorosis, recent evidence suggests there may also a genetic component. Research by Everett, McHenry, Reynolds, Eggertsson, Sullivan, Kantmann, Martinez- Mier, Warrick, & Stookey (2002) found that in mice there appears to be genetic variation in susceptibility to fluorosis. However there is currently insufficient evidence to begin making predictions about the genetic influence on fluorosis in humans.

1.3.3. Prevalence of fluorosis

The best evidence from a recent, systematic review of the prevalence of fluorosis shows that at the optimal fluoridation level, 48% of people drinking the water will have some degree of fluorosis, and 12.5% of people will have fluorosis that is at a level which causes aesthetic concern (McDonagh et al., 2000). However there is a great deal of variation among individual papers both in the actual level of fluorosis reported, and the type of information reported (some report the prevalence of all fluorosis, some the prevalence of fluorosis which is an aesthetic concern, some report whether or not the prevalence has increased).

The prevalence of fluorosis is most commonly reported in the United States of America. Clark's (1994) review found that the prevalence of dental fluorosis in the U.S.A. had increased since the time of Dean. He estimated the prevalence was between 35-60% in fluoridated areas and between 20-45% in non-fluoridated areas, depending on local conditions. He also found some evidence that the prevalence of severe and moderate fluorosis is increasing, not just mild fluorosis. Beltran-Aguilar, Griffin, & Lockwood (2002) found that the prevalence of fluorosis in areas of optimal fluoridation (0.7-1.2 ppm F^-) was 25.8%. Where the fluoridation level was higher than optimal (0.7-4.0 ppm F^-) the prevalence was 37.8%, and where the fluoridation level was sub-optimal (<0.7 ppm F^-) the prevalence was 15.5%. Like Clark (1994), they noted that the prevalence of fluorosis had increased between the 1930's and 1980's in the U.S.A., with the largest increase occurring in the sub-optimal fluoridation sample (6.5%-15.5%).

A literature review by Szpunnar & Burt (1987) also reported a slight increase in the prevalence of fluorosis in a number of communities when compared to the prevalence at the time of Dean. However they acknowledged this apparent increase was only slight and is at least partially accounted for by different examining conditions. They also note that it is possible that aesthetic standards are higher now than they were in the Depression-era. Although they acknowledge their review is not conclusive they believe that dental fluorosis should continue to be monitored, especially in light of the numerous sources of fluoride currently available. However, Lalumandier & Rozier (1998) reported a much higher prevalence of fluorosis, in their sample of American children it was almost 78%, which was six times higher than the prevalence in the late-1930's and three times higher than that observed in the mid-1980's.

In a subsequent review of the literature of the last 50 years, Rozier (1999) found that the prevalence of fluorosis had increased in communities where the fluoride in the drinking water was less than 0.3 ppm. He also noted that an increase in the prevalence of fluorosis in those drinking optimally fluoridated water has occurred. In his review Rozier (1999) cited the 1986-87 national survey of U.S. school children conducted by the NIDR (which included assessments of dental fluorosis) which he considered to be of importance because it establishes "a national baseline for future comparisons, as well as current reference for regional, state, or local surveys" (p.241). The NIDR survey found 78% of the total U.S. population 7 years of age or older did not have any definite signs of fluorosis, and of the 22% who did have definite fluorosis 76.2% of these were very mild fluorosis. In other words 5.28% of Americans 7 years old or above have fluorosis at a level above very mild.

The fact that both Beltran-Aguilar et al. (2002) and the Rozier (1999) review reported an increase in the prevalence of fluorosis in areas of sub-optimal fluoridation, could be taken as evidence that fluoride supplements are leading to an increase in fluorosis, since there is no reason to believe the water supply is responsible for these increases. The issue of fluoride sources other than that found in the drinking water was discussed under the aetiology of fluorosis (Section 1.3.2).

Griffin, Beltran, Lockwood, & Barker (2002) compared estimates of fluorosis prevalence and risk attributable to water fluoridation using an index applied to the entire dentition (Dean's index; Dean, 1934) and an index applied to the maxillary anterior teeth (anterior fluorosis index, which they developed themselves). They also estimated the aesthetic problems that were attributable to fluoridation. They found that approximately 2% of U.S. school children might experience perceived aesthetic problems which could be attributed to the recommended levels of fluoride in the drinking water. The total prevalence of fluorosis was 26% (using Dean's index, or 18% using the anterior fluorosis index), however they believed only 2% of these participants had fluorosis that was an aesthetic concern that was attributable to water fluoridation. These findings are clearly much closer to the work of Beltran-Aguilar et al. (2002) than of Lalumandier & Rozier (1998). They also found that using the anterior fluorosis index yielded a lower prevalence of fluorosis than was found with Dean's index. Using Dean's index 33% of children in the optimally fluoridated group and 9% of children in the low fluoridation group had fluorosis that was very mild or greater. Whereas when using the anterior fluorosis index, 23% of children in the optimally fluoridated group and 5% of children in the low fluoridation group had fluorosis that was very mild or greater. Therefore it may be the case that some of the reported fluorosis effects teeth that are not easily viewed and therefore has a small aesthetic impact.

While most of the above reviews were international in their scope (Bowen, 2002; Horowitz, 1992; Levy et al., 1995; McDonagh et al., 2000; Newbrun & Horowitz, 1999), the individual studies described have all been concerned with the prevalence of fluorosis in the U.S.A. Outside the U.S.A., Whelton, Ketley, McSweeney, & O'Mullane (2004) found that the prevalence of fluorosis may be increasing in the European Union, but it was difficult for them to draw any firm conclusions due to the lack of a standardised method of measurement in the studies they reviewed. In Australia, Riordan (2002) found that the prevalence of dental fluorosis declined after the School Dental Services in Western Australia chose to initiate changes in the fluoride-supplement and toothpaste regimens of the children of that region. For example they discouraged the use of fluoride supplements by children and recommended children use low fluoride toothpaste. In 1989-1990 the prevalence of fluorosis was 40.2% in the Perth sample (a fluoridated area, at the level of 0.8mg/L) and 33.0% in the Bunbury region sample (which is an unfluoridated area, the fluoride level is approximately 0.25mg/L). In 2000 the prevalence of fluorosis had fallen to 22.2% in the Perth sample and 10.8% in the Bunbury region sample. Furthermore no increase in dental caries was recorded in these samples.

Riordan believed these results suggested that fluorosis could be "brought under control at population level by eliminating discretionary fluoride intake from supplements and reducing the possibility of intake from toothpaste, without apparent adverse changes in the impact of dental caries" (p.239). This further corroborates the reviews above which proposed that fluoride supplements could explain the increased prevalence of fluorosis.

Holloway & Ellwood (1997) reviewed studies of fluorosis in Britain. They noted that there was a lack of in-depth studies and acknowledged the conclusions they drew are tentative and inconclusive. They reported "no substantial evidence of increases in either DDE (Developmental Defects of Enamel) or fluorosis" (p.148). However they also say that from more recent reports it appears that:

> The prevalence of developmental defects of enamel in communities receiving optimally fluoridated water is now slightly higher than that in communities receiving fluoride-deficient water and that this may not have been the case some 20 years ago (p.151-2).

They qualified this by pointing out that these findings may be due to the prevalence of diffuse opacities or by the use of a more sensitive index.

Overall it is unclear what the prevalence of fluorosis is, but perhaps it is wise give most consideration to the detailed review conducted by McDonagh et al. (2000), that suggested we should expect to see some level of fluorosis in almost 50% of people in optimally fluoridated regions, but that fluorosis will only be an aesthetic problem for around 13% of the population. While studies in the U.S.A. seem to show that the prevalence of fluorosis is increasing, they have not agreed at what rate this is happening, but there seems to be mounting evidence that the prevalence of fluorosis can be reduced by the careful management of fluoride supplements.

1.3.4. Perceptions of fluorosis

There are numerous ways to consider the perception of fluorosis. It can be considered from the viewpoint of: the affected individual; the parents of affected children; the dental professional; and finally it can be examined from the general public's perspective.

1.3.4.1. The perspective of affected patients and parents of affected children

Numerous studies that have considered the affected patients' (or their parents') perception of fluorosis have found that fluorosis can be the source of dissatisfaction. Several studies have looked at the attitudes of individuals affected with severe fluorosis. These studies, which have tended to be conducted in Africa, or with participants of African origin, (since there is little severe fluorosis in Europe or America), appear unanimous in their finding that severe fluorosis is perceived as a problem in affected individuals. van Palenstein Helderman and Mkasabuni (1993) reported that individuals in Tanzania who were affected with severe fluorosis were unhappy with their appearance and had a desire for cosmetic treatment.

Rodd & Davidson (1997) noted severe fluorosis can be seen in people from areas where there is a very high level of naturally occurring fluoride in the water, and gave the example of Somalia. They also claimed that "discoloured enamel is accepted as normal (or even attractive) in the native country" (p.409) but not in the U.K., and young Somali refugees were increasingly requesting cosmetic treatment for their anterior teeth. However these statements did not appear to be based on any specific study or reference. They also reported a case study of an 11-year old boy who was teased at school because of his discoloured teeth; his dental practitioner subsequently referred him for management of his fluorosis. This teasing is one example of the way in which dental fluorosis could cause psychological distress.

Numerous studies have attempted to identify the level at which fluorosis becomes a cause for concern in affected individuals and their parents. Chikte, Louw, & Stander (2001) investigated the perception of fluorosis in communities living in the Northern Cape Province of South Africa. There is a considerable range in fluoride levels in the drinking water in this area and they reported that concerns with tooth appearance were mostly –but not exclusively- expressed by children with moderate or severe fluorosis. Studies that assess opinions regarding the milder forms of fluorosis are inconclusive. Lalumandier & Rozier (1998) found that parents of children without fluorosis were two and a half times more likely to be satisfied with the colour of their children's teeth than parents whose children had fluorosis. Even when considering very mild fluorosis they found that a significantly greater proportion of parents whose children had very mild fluorosis were dissatisfied with their children's appearance when compared with parents whose children had no fluorosis. Moreover, parents' level of satisfaction with the colour of their children's teeth decreased as the severity of the fluorosis increased. Lalumandier & Rozier (1998) noted that the relationship between enamel fluorosis and psychological distress is unknown, and called for research to "determine what effect fluorosis has on dental aesthetics while controlling for other non-fluoride related opacities" (p.1005).

In contrast to Lalumandier & Rozier's (1998) findings not all studies have found mild fluorosis to be of aesthetic concern to parents. Bowen (2002) reviewed the literature concerning fluorosis and reached the conclusion that mild fluorosis was not a cause for concern; he felt "the mottling generally is not noticed by most people or, indeed, by the individual himself or herself" (p.1047). Further to this, Woodward, Main, & Leake (1996) investigated parents' satisfaction with their child's teeth in Ontario, Canada. In this study of 385 parents whose 8-9 year old children had undergone a dental examination, 114 (29.6%) were not satisfied with the appearance of their children's teeth. Although the scores for five measures of oral health were obtained while examining the children (malocclusion, fluorosis, periodontal health, calculus, and dental caries) only fluorosis and malocclusion showed significant associations with parental satisfaction. The parents of children with a TSIF score (Tooth Surface Index of Fluorosis: Horowitz, Driscoll, Meyers, Heifetz, & Kingman, 1984) of 2 or greater were half as likely to be satisfied with the appearance of their child's teeth as the parents of a child with no or mild fluorosis (TSIF of 0 or 1).

Clark, Hann, Williamson, & Berkowitz (1993) investigated the attitudes of children with fluorosis and their parents to the child's tooth appearance. They reported that 60% of their sample of school children had dental fluorosis on at least two tooth surfaces. Of these 52% were classified as having a score of 1 on the TSIF, and the remaining 8% had scores between 2 and 6. Parental and child ratings of satisfaction with the childrens' tooth colour were taken, although the source of any dissatisfaction was not assessed. The results showed that generally speaking parents had more concerns about the tooth colour than their children, but of the children with a TSIF score of 1 very few were regarded has having a problem by themselves or by their parents. The number of children and parents who were concerned with the childrens' tooth colour increased when the TSIF score was between 2-6 but even for scores of 2 or 3 few parents or children perceived an aesthetic problem.

Sigurjons, Cochran, Ketley, Holbrook, Lennon, & O'Mullane (2004) investigated parental perceptions of fluorosis in Iceland, Ireland, and England. They found that the higher the TF score the children had, the more concern there was about their teeth, but only at a TF score of 3 (moderate fluorosis) did fluorosis become an appreciable concern. Therefore, parents' perception of very mild fluorosis varies across studies, and requires further study to find the cause of this ambiguity.

Recent work has begun to consider the possibility that fluorosis may impact on affected individuals and their parents in ways that go beyond dissatisfaction with the aesthetic appearance of teeth. Two abstracts published in the Journal of Dental Research found that children's perceptions of fluorosis appeared to have an impact on their well being (Soto-Rojas, Martinez- Mier, Maupome, & Urena-Cirett, 2002), and that parents perceptions' of fluorosis in their children's teeth appeared to have an impact on their well being (Urena-Cirett, Martinez- Mier, Maupome, & Soto-Rojas, 2002).

1.3.4.2. The perspective of dental professionals

Studies have also been conducted on dental professionals' perceptions of fluorosis. Levy, Warren, & Jakobsen (2002) asked fourth year dental students to repeat a study they had participated in as incoming first year students. The study assessed their perceptions of fluorosis by asking them to compare photographs of fluorosis to various other conditions. The photographs had been created using computer technology to impose the dental conditions on a standardised, computer generated dentition (McKnight, Levy, Cooper, Jakobsen, & Warren, 1999). The "fourth-year dental students generally had more favourable aesthetic perceptions of mild dental fluorosis and other conditions than they had reported as entering students" (p.24). This study suggested that there may be an adaptive effect in dental professionals leading to them becoming more tolerant of the aesthetic effects of mild fluorosis than lay people.

Clark (1995) asked pairs of children and parents, and also dental professionals to assess the dental aesthetics of fluorosis from the different categories of the TSIF, as displayed on slides. Some of the children in the sample were the subjects of the photographs that were used. Results showed that parents, children and dental professionals could distinguish between teeth with fluorosis and those with no fluorosis, and that teeth with no fluorosis were preferred. In addition to this they also found that for low levels of fluorosis (TSIF of 0 and 1) or for non-fluorotic defects (TSIF=8), dental professionals rated aesthetics significantly better than parents who, in turn, rated significantly better than children. This adds weight to the concerns of Riordan (1993b), who raised the possibility that parents and lay people may be more concerned than dentists about the aesthetic aspects of dental fluorosis. However it is also worth noting that the raters showed poor internal validity for duplicated slides, with parents showing the poorest consistency and dental professionals showing the best consistency. Clark (1995) also noted that there were dramatic differences in raters' opinions of non-fluorosed teeth and teeth with a TSIF score of 2-6. This suggests that fluorosis may be noticed in its milder forms but it only becomes a cause for aesthetic concern when it is at the moderate or severe level.

However, not all studies have found dentists to be more tolerant of the aesthetic impact of fluorosis than lay people. Milsom, Tickle, Jenner, & Peers (2000) investigated perceptions of developmental defects of enamel fluorosis in 12-year-old children in Crewe, U.K. They found that while a dental epidemiologist identified enamel defects in the upper incisors of 34% of the children, less than 7% of the sample reported they were unhappy with the appearance of their teeth because of "marks that would not brush off". Moreover, one in 20 of the children identified by the dental professional as having no developmental defects of enamel on their upper incisors reported being unhappy with the appearance of their teeth because of "marks that would not brush off". The sample used by Milsom et al. (2000) contained children from both fluoridated and non-fluoridated communities, no statistically significant differences in the prevalence of enamel defects were found between these communities. This was true for both normative and subjective judgements.

While the perceptions of dental professionals and affected patients can be informative in assessing fluorosis, the perspective of most importance to this thesis is that of the general public. This aspect will now be described and discussed.

1.3.4.3. The perspective of the general public

Riordan (1993a) reported that when lay people were shown children who had fluorosis, they found fluorosed teeth (with a TF score of 2 or greater) less aesthetically pleasing than teeth with no fluorosis or with mild fluorosis. Lay people also mistook fluorosis as a sign of neglect of dental hygiene.

Ellwood & O'Mullane (1995) considered the importance of demarcated opacities and enamel hypomineralisation (fluorosis) on dental aesthetics from several perspectives. They asked lay observers to rate the dental aesthetics of 21 patients in an actual meeting as opposed to viewing photographs of them. They found that lay observers gave people with mild fluorosis a similar aesthetic rating to people with no fluorosis, but individuals with moderate or severe fluorosis were regarded as having a significantly worse dental appearance than those with no fluorosis or mild fluorosis. These results were mirrored by the patients themselves and by dental professionals.

Other studies that found mild fluorosis not to be of aesthetic concern include Alkhatib, Holt, & Bedi (2004) and Hawley, Ellwood, & Davies (1996). The latter conducted a study with 14-year old adolescents and found that when shown photographs of various levels of fluorosis they rated scores lower than 3 on the TF scale as being of little aesthetic concern, in fact mild fluorosis was more acceptable in appearance than no fluorosis. In a cross-sectional national survey in the U.K., Alkhatib et al. (2004) found that only 14% of respondents thought that mild fluorosis was aesthetically objectionable, rising to 45% for moderate fluorosis and 91% for severe fluorosis.

However, McKnight, Levy, Cooper, & Jakobsen (1998) and McKnight et al. (1999) found that even mild fluorosis could be viewed as an aesthetic concern. McKnight et al. (1998) compared different levels of fluorosis with various other dental conditions (normal teeth versus mild fluorosis, open bite versus moderate fluorosis, rotated teeth versus mild fluorosis, tetracycline staining versus moderate, isolated opacities versus severe fluorosis) in each comparison the fluorotic teeth scored less favourably than the non-fluorotic teeth. McKnight et al. (1999) compared aesthetic perceptions of mild fluorosis and other conditions using computer-generated images. The standardised images allowed a better comparison of fluorosis with the other conditions. The comparisons were; normal teeth versus mild fluorosis, diastema versus mild fluorosis, isolated opacity versus more involved mild fluorosis, and mild fluorosis on the incisal third only versus more generalised mild fluorosis. With the exception of diastema versus mild fluorosis, non-fluorotic teeth were preferred in each comparison (in the case of mild fluorosis on the incisal third only, versus more generalised mild fluorosis, the more generalised mild fluorosis was preferred).

It is apparent from the literature that previous work has not yet conclusively determined the level at which fluorosis becomes an aesthetic problem, and has not discovered how fluorosis is perceived by the general public. Whelton et al. (2004) have called for further research into the aesthetic impact of fluorosis, and Chikte et al. (2001) argued a case for an assessment of the public's perception of fluorosis and societal dental norms. It appears that when considering fluorosis from the point of view of the affected patient or the parent of an affected child, assessments of the perception fluorosis are moving beyond just its aesthetic impact (Soto-Rojas et al., 2002; Urena-Cirett et al., 2002). Furthermore, studies of other dental conditions have also considered whether or not the public make attributions that go beyond the aesthetic in response to dental appearance (Eli et al., 2001; Feng et al., 2001; Newton et al., 2003). This approach may prove beneficial when considering the public's attitudes to fluorosis. Determining what assumptions the public make about individuals with fluorosis, and what personality traits they attribute to them (if any) will make future investigations of, and decisions about, water fluoridation better informed. For example, there is some evidence that fluorosis may be mistaken for a sign of dental neglect (Riordan, 1993a). If people believe fluorosis is caused by dental neglect do they assume those who have fluorosis are lazy or unhygienic? It is also necessary to establish at what severity of fluorosis such assumptions are made. If the public does not notice mild fluorosis, or if they do notice it but they do not make negative assumptions about it, then should the fluoridation policy be altered for a community where mild fluorosis is relatively common but the more severe levels of fluorosis are rare?

There is also the issue of balancing the costs of fluorosis against the benefits of reduced caries. While considering the social impact of fluorosis could be informative for public health debates about water fluoridation, this needs to be balanced against the social consequences of dental caries. If fluorosis is viewed as the cost of fluoridation, then the benefit is reduced caries. Like fluorosis, caries may be detrimental to facial appearance, but unlike mild fluorosis they also require dental treatment and can involve physical pain. Therefore a comparison of the fluorosis and caries in terms of the impact they have on appearance would be of value.

1.3.5. Fluorosis and the current research

The literature discussed up to this point has addressed the importance of appearance, and in particular dental appearance, and how deviating from a normal appearance can be the cause of negative social consequences and emotional upset. It has also discussed the aesthetic impact of dental fluorosis and highlighted gaps in the fluorosis literature that, were they filled, might provide guidance on future policies regarding both water fluoridation and other fluoride sources such as tablets and toothpastes. The final section of the literature review will address the best way to identify what social impact fluorosis has. In order to achieve this it will consider the most effective way of measuring attitudes, and then link this to the measurement of the public's attitudes to fluorosis.

1.4. The measurement of attitudes

1.4.1. Introduction

The definition of attitudes is not straightforward. Reber (1995), for example, noted that "exactly how the term (attitude) is used in modern psychological literature will... depend largely on the theoretical tilt of the writer". For the purposes of this thesis Fazio, Chen, McDonel, & Sherman's (1982) definition of an attitude as being simply an "association between a given object and a given evaluative category" (p.341) will be used, since the work of Fazio and his colleagues is prominent in attitude measurement.

Measures of attitude have traditionally been conducted using self-reported or explicit scales. The most common of these are the Thurstone scale (Thurstone, 1927a; Thurstone, 1927b; Thurstone, 1928), the Guttman scale (Guttman, 1944), the Likert Scale (Likert, 1932), the semantic differential scale (Osgood, Suci, & Tannenbaum, 1957), the visual analogue scale, and the feeling thermometer, a variation of the visual analogue scale, created because it is easier to use. From reading the past literature it is unclear who invented the latter two scales but they have been commonly used in health psychology and measuring political opinion. These scales share many similar properties and a detailed analysis of these measurement techniques is beyond the scope of this thesis (but can be found elsewhere, Himmelfarb, 1993; Moser & Kalton, 1972; Oppenheim, 1992; Streiner & Norman, 1995). However, they will now be briefly summarised.

In order to create a Thurstone scale the experimenter generates many statements about a particular attitude object (perhaps between 100-200) and writes each one on a card. Then he or she asks judges to place each statement into a set number of piles of increasing favourableness. The experimenter then selects the statement from each pile that shows the greatest consensus and using these statements, creates a scale of increasing favourability in which each statement is assigned a numerical value to indicate its position on the scale. The Thurstone scale assumes that the statements are normally distributed and in this way attempts to generate interval data. Once the scale is created it is administered to the participants, whose score is calculated as the mean or median of the scale values of each item they endorsed.

The Guttman method is a scale that produces ordinal data. Like the Thurstone scale it begins with a large number of statements, but this number is reduced by the experimenter to a relatively small number of items (10-20). Unlike the Thurstone scale these items are not sorted into intervals by judges but by the respondents themselves and it is crucial that the statements all relate to a single dimension of the attitude being measured. The responses are put into a respondent-by-item matrix showing whether each participant responded "yes" or "no" to each item ("yes" may be represented by 1 and "no" by 0). This matrix is known as a scalogram, and this type of analysis is sometimes referred to as scalogram analysis. The Guttman method relies on cumulative properties, that is the assumption that each respondent will endorse all of the items that are less extreme than the most extreme item he or she has endorsed. In reality this is quite rare and so Guttman proposed a coefficient of reproducibility, which measures the extent to which the respondents' endorsements can be reproduced from the triangular relationship that defines a perfect Guttman scale. This coefficient was not as informative as intended and alternative measures have been proposed, but the difficulty in using Guttman scales has meant that easier to use methods are often preferred (Himmelfarb, 1993).

Both the Thurstone scale and, in particular, the Guttman scale can be cumbersome to use. The Likert scale provides a simpler alternative. Likert scales consist of a series of statements about an attitude-object; each one is typically followed by a five-point scale (although other numbers have been used). Each point of the scale has a label assigned to it indicating one level of agreement or disagreement with the statement, and the middle point usually indicates neutrality. In a conventional Likert scale the five points would be “strongly disagree”, “disagree”, “neutral/don't know”, “agree”, and “strongly agree”. Participants respond by indicating their level of agreement to each statement, and then a total score is calculated. Some of the items may need to have their score reversed so that the number five always indicates a strongly favourable opinion to the attitude object, and the number one always indicates a strongly unfavourable opinion. The statements used would usually be selected intuitively by the experimenter and then the most suitable would be identified by pilot testing.

Semantic differential scales are very simple to design. They typically consist of seven point scales anchored by polar opposite adjectives. The attitude object being measured is written at the top of the page and then participants are asked to rate this object by selecting a point on each of the scales they feel best describes it. The scales are often subject to factor analysis to determine if they are interrelated. The adjectives used are very general and have clear evaluative meaning so items do not need to be prepared in advance and scaled; this makes the semantic differential scale the easiest to use of the measures described thus far.

Visual analogue scales are similar to the semantic differential scale in that they are anchored by polar opposite descriptions. They are different in that they have no intermediate positions between the anchors; instead there is simply a straight line of 100mm on which participants are asked to mark their attitudinal position. The experimenter measures where the mark was placed and this gives the respondents score out of 100. The visual analogue scale is not mentioned in many books of attitude measurement (Himmelfarb, 1993; Moser & Kalton, 1972; Oppenheim, 1992). However, it appears to be more popular in health measurement scales, where, for example, participants may be asked to indicate their level of pain on a visual analogue scale anchored with no pain, and extreme pain (Streiner & Norman, 1995). Although the visual analogue scale is relatively easy to use it may be unsatisfactory to elderly respondents who prefer vertical lines, a problem which led to the development of feeling thermometer (Streiner & Norman, 1995). The feeling thermometer ranges from 0-99 degrees, and asks participants to rate how “warm” or “cool” they feel towards an attitude object by indicating which “temperature” reflects their attitude on an illustration of a thermometer (Greenwald, McGhee, & Schwartz, 1998).

While the Thurstone scale aims to provide interval data and involves steps to try and ensure this (although the validity of this has been questioned by Moser & Kalton, 1972), and the Guttman scale produces ordinal data, there is some debate surrounding whether Likert scales and semantic differential scales produce ordinal or interval data. Himmelfarb (1993) believed it is unclear what type of data they produce, and Streiner & Norman (1995) noted that while technically they give ordinal data it may be acceptable to treat it as interval data. However, Moser & Kalton (1972) insisted that they clearly give ordinal data. The type of data obtained by visual analogue scales and feeling thermometers has not been debated to the same extent as Likert scales and semantic differential scales. While it is clear that the difference between any two points on these scales is just as arbitrary as the difference between two points on a Likert/semantic differential scale, there are no labels on these scales with which to categorize responses. However, since the data from feeling thermometers, visual analogue, Likert, and semantic differential scales are conventionally analysed as if they are interval data this is a theoretical question rather than a practical issue and has not detracted from the use of these scales.

Each of the scales has unique advantages and disadvantages but they often share some common problems such as bias in responding. There are several types of bias and some will not apply to every type of scale or be relevant to every area of attitude measurement. However, a brief overview of some of the issues surrounding self-report measures will be described to outline why they are not always the best option for attitude measurement. Problems with self-report measures are described below.

Differences between the creators of, and responders to, a scale

Bias can be caused by differences in perspective between the creators of and responders to a scale (Streiner & Norman, 1995). An example of this was provided by Rutt & Coleman (2001), who found that two questionnaires that were designed to measure body image, the Body Image Questionnaire (Berscheid, Walster, & Bohrnstedt, 1972) and the Eating Attitudes Test (Garner & Garfinkel, 1979) using white American participants were more effective for use on a Hispanic population after being redeveloped within that population.

Differences between respondents in their reference groups

In addition to differences between the researchers and participants there may be variation between participants' responses caused by having different reference groups. For example, Tanzer (1996) reported there are cultural differences in how willing people are to self-praise, with Australian students more willing to self-praise than Singaporean students. Furthermore, sometimes self-report measures may fail to identify cultural differences between respondents when they exist because respondents only compare themselves to people from the same culture. Heine, Lehman, Peng, &

Greenholtz (2002) found that self-report scales did not identify any difference between North American and East Asian participants in how individualistic-collectivistic they were, despite agreement among cultural experts that East Asians were more collectivistic. There can also be different reference groups within a culture. Marsh, Hey, Roche, & Perry (1997) found that students who were not elite athletes but attended a specialist sports school had a lower physical self-concept score than students who were not elite athlete students but attended a typical school. They attributed this to the fact that the students at the normal school were not comparing themselves to elite athletes.

Optimizing and satisficing

Optimizing and satisficing refer to whether participants answer the questions to the best of their ability (optimizing), or whether they respond in a way that while satisfactory to the investigator requires as little effort as possible from them and probably does not reflect their true feelings, which is known as satisficing (Streiner & Norman, 1995).

Other types of bias

Streiner & Norman (1995) also noted several other types of bias including: positive skew (the tendency to give more favourable responses on ordinal or interval scales); halo bias (the tendency for participants overall impression of the attitude object they are rating to influence their individual responses); and the central tendency bias (the tendency not to endorse the extremes of a scale but to stay nearer the middle). They also warned that the wording or framing of the questions can cause bias by leading respondents to a certain answer.

Further to these general concerns, two types of bias are of particular interest to social psychology in attitude measurement. These are social desirability concerns and respondents' lack of awareness of their own attitudes.

Social desirability refers to participants responding in a way that is intended to portray them in favourable light (or in some cases an unfavourable light, e.g. when they wish to qualify for a health programme and so try to impress staff with the severity of their condition; Streiner & Norman, 1995). The exact motivation for, and manifestation of, respondents' socially desirable answers will depend on the attitude being measured. For instance, several studies have been conducted that show a discrepancy between self-reported racial prejudice and implicitly measured racial prejudice (Sections 1.4.2 and 1.4.3, also see Fazio & Hilden, 2001; Fazio, Jackson, Dunton, & Williams, 1995; Phelps, O'Connor, Cunningham, Funayama, Gatenby, Gore, & Banaji, 2000), a similar trend was found in religious, age, and nationality prejudices (Rudman, Greenwald, Mellott, & Schwartz, 1999).

Whether or nor participants can be unaware of their own attitudes is difficult to determine, but some authors believe we can hold attitudes that we are introspectively unaware of, and these attitudes may be measured using implicit techniques (Greenwald & Banaji, 1995). Other researchers argue there is no evidence that people are unaware of their attitudes, and it is only the measure that is implicit, since although participants are unaware that their attitudes are being measured they are not necessarily unaware that they hold those attitudes (Fazio & Olson, 2003). What has been shown is that implicit measures can explain variance not accounted for by explicit measures in numerous fields, including shy behaviour, racial prejudice, self-esteem, and self-concept (Asendorpf, Banse, & Mucke, 2002; Cunningham, Preacher, & Banaji, 2001; de Jong, 2002; Greenwald & Farnham, 2000; Greenwald, McGhee, & Schwartz, 1998). Whether this is attributable to participants deliberately not reporting their attitudes or not reporting them through lack of awareness is still a matter for debate (although it seems participants would have little motivation to hide their true attitudes on some of the implicit measures that have diverged from explicit measures, such as preferring flowers over insects, Karpinski & Hilton, 2001). Nevertheless, if implicit measures can uncover more of our attitudes than would be possible with explicit measures alone there is clearly a case for their use.

De Houwer (2003) summed up the main concerns with self-report measures when he said "direct measures of attitudes are susceptible to deception and self presentational strategies... (and that furthermore) in daily life people often do not analyse their attitudes towards objects in a conscious and deliberate manner". He continued "Most authors agree that direct measures of attitudes are not well suited to measure the spontaneous affective reactions to attitude objects that often guide behaviour." (p.219). In other words there are two main limitations to self-report measures, the first is that people may be unwilling to report their true attitudes: the second is that they may be unable to report them.

These concerns have led to the increasing use of implicit measures of attitude (De Houwer, 2003). Musch & Klauer (2003) propose that although evaluative and affective information processing have long been a fundamental issue in social and cognitive psychology, there has been a dramatic increase in their interest and evaluative processes is now one of the most rapidly growing endeavours of psychology.

This is not to imply that self-report measures are redundant in measuring attitudes. It has been shown that explicit measures of attitude can sometimes predict behaviour more successfully than implicit measures, especially for deliberative or controlled behaviour e.g. choice of snack when offered an apple or a candy bar and condom use with a steady partner (Karpinski & Hilton, 2001; Marsh, Johnson, & Scott-Sheldon, 2001; this is discussed further in Section 1.4.3.3). Furthermore, it may be beneficial to use both types of measure in a clinical setting (de Jong, 2002).

1.4.1.2. Implications for the present study

Previous work studying attitudes to fluorosis and to dental appearance in general has relied entirely on explicit measures (Sections 1.2.5 and 1.3.4). While this work has not found participants to be reluctant in expressing their opinions, even when these are negative, one of the objectives of this thesis is to investigate if people attribute personal characteristics to individuals affected with fluorosis that go beyond aesthetic preferences. Participants' responses may be guarded when answering such questions due to social desirability concerns. That is, participants may not wish to respond in a way that shows them to make judgments about people based solely on their dental appearance, for fear of seeming rude or superficial. This may lead them to answer questions in a way which understates the impact dental fluorosis has on a persons appearance. Therefore, alternative ways of measuring attitudes which make it difficult to consciously modify responses, such as implicit measures, may be valuable in assessing attitudes to fluorosis.

1.4.1.3. Implicit measures of attitude

Unlike explicit measures of attitude, where the respondent is aware that their attitude is being assessed, implicit measures of attitude refer to the measurement of attitude without the respondent's awareness. The most common methods of achieving this involve using procedures known as the Affective Priming Task (APT) and Implicit Association Test (IAT) (Fazio & Olson, 2003). De Houwer (2003, p.220) stated that the APT is the "first and most influential" of the reaction time tasks used to indirectly measure attitude. Steffens & Plewe (2001) proposed that just two years after it had been developed the IAT had already changed the field of social cognition in a profound way.

Research involving the APT and IAT will be reviewed separately and then compared with each other in terms of their suitability for use in the current investigation.

1.4.2. The Affective Priming Task (APT)

The APT was introduced by Fazio, Sanbonmatsu, Powell, & Kardes (1986). They based it on work done in the area of priming in spreading activation processes (e.g. Neely, 1977). This work focused on the latency of decisions regarding whether a string of letters was a word or non-word, and on how this decision was facilitated by the prior presentation of a prime. Evidence from this type of experiment showed that lexical decisions were facilitated when respondents were exposed to a prime that was related to the word being judged. When the word being judged as word or non-word was "nurse", for example, it would be facilitated by the prime "doctor" because they are conceptually linked and therefore the word "nurse" is activated upon seeing the word doctor (Fazio, 2001). Fazio et al. (1986) postulated that attitudes may be activated through a similar associative effect when the perceiver is primed by a relevant stimulus.

This priming effect was demonstrated when Fazio et al. (1986) showed that participants were able to evaluate adjectives more quickly when these adjectives were immediately preceded by (i.e. primed by) attitude objects of a similar valence, compared with when these adjectives were primed by attitude objects of an opposite valence. In their study, participants viewed a monitor screen and used a response box with two keys on it. A series of words appeared on the screen consecutively, the words were paired so that one trial consisted of two consecutive words. Participants were told to hold the first word that appeared on the screen (i.e. the prime) in memory and to say it out loud after the second word disappeared. When the second word (i.e. the target word) came onto the screen, they were instructed to press the first key if it was positive in meaning, and to press the second key if its meaning was negative. Their response times proved faster when the prime word and the target word were of the same valence: that is when the target word was "disgusting" responses were faster when the prime was "cockroach" than when the prime was "butterfly". This facilitation effect occurred when the prime and target matched in valence despite the apparent irrelevance of the prime to the task.

There were, however, limits to their findings. Fazio et al. (1986) only found the facilitation effect when the time between the prime and target word was very short. The effect was found if the time difference -or stimulus onset asynchrony (SOA) - was 300ms but not when it was 1000ms. This finding has since been replicated (e.g. De Houwer, Hermans, & Eelen, 1998) and is important as Fazio et al. (1986) contended that if the facilitating effect of the prime was a conscious, effortful, process then it would be expected that a longer SOA would strengthen the effect because it would give participants longer to evaluate the prime. However the effect disappeared with a longer SOA, suggesting the activation occurred automatically and dissipated very quickly or else had been suppressed because it was irrelevant to the task the participant was performing.

1.4.2.1. The generality of the affective priming effect

The priming effect has been reproduced in a number of different experiments since Fazio et al. (1986) and has proven to be a robust phenomenon that occurs across a range of prime stimuli, target stimuli, and specific task requirements (Fazio, 2001). Affective priming has been demonstrated with various primes, including the names of familiar attitude objects (Bargh, Chaiken, Govender, & Pratto, 1992), colour images (Fazio, 1993b), black-and-white line drawings (Giner-Sorolla, Garcia, & Bargh, 1999), Turkish words that effectively served as non-words, (De Houwer et al., 1998) and even odours (Hermans, Baeyens, & Eelen, 1998). Nouns and colour images have also been used as target stimuli (Bargh, Chaiken, Raymond, & Hymes, 1996; Hermans, De Houwer, & Eelen, 1994).

The affective priming effect has also been found under different task requirements, including participants being asked to keep the prime word in memory (Fazio et al., 1986), to study the prime so that it can be identified at a later time (Fazio, 1995), and not being given any specific task (Bargh et al., 1992). Furthermore, Bargh et al. (1996) demonstrated the affective priming effect by asking participants to ignore the prime word and say the target word aloud as quickly as possible, meaning participants were not deliberately evaluating anything. Yet they were still influenced by the prime. The effect has also been found when the prime was presented subliminally (e.g. Niedenthal, 1990, Wittenbrink, Judd, & Park, 1997) providing further support for the idea that the effect is an automatic process.

Studies have also used word recognition tasks rather than tasks of valence judgements. Fazio, Williams, & Powell, (2000), for example used a design in which the target stimuli was masked by a block of dots which gradually disappeared until the word was revealed. Participants were asked to name the word as soon as they recognised it. Recognition occurred more quickly if the target word was primed by a word of similar valence to the target word. Duckworth, Bargh, Garcia, & Chaiken (2002) and Garcia & Bargh (2003) found that the affective priming effect can even occur following the presentation of visual and auditory novel stimuli including altered abstract art and Turkish words.

In another variation of the APT, Greenwald, Draine, & Abrams (1996) introduced a response window innovation, which forced participants to respond in a certain time period after the presentation of the target. The dependent variable is the number of errors participants make in classifying the target rather than the length of time they take to classify the target. The response window ensures the priming affect cannot be diluted by being spread over two variables, namely response latency and response accuracy, but is instead concentrated on response accuracy. Klauer & Musch (2003) note the recent success of the response window variation in a number of studies.

1.4.2.2. Mechanisms mediating affective priming

This section describes three models that have been used to explain the affective priming effect; the spreading activation mechanism, the affective-matching mechanism, and the response competition mechanism. The last of these models accounts for more of the research than the previous two, but it seems unlikely any of the models in isolation can fully explain the effect, and rather a combination of the models may be the key to developing further understanding of affective priming.

Spreading activation mechanism

Bargh et al. (1996) proposed a spreading activation model to explain affective priming, which consisted of a "semantic associative network" (p.121). They believed the explanation for one positive word (such as "flowers") being able to facilitate the pronunciation of another positive word (such as "knowledge") despite having no connection to that word beyond a shared valence, was that all positive concepts were interconnected to each other in memory and all negative concepts were interconnected to each other. When one positive concept was activated (by serving as a prime) this activation spread throughout the positive network activating all other positive concepts. However, this model has been rejected by most researchers (Klauer & Musch, 2003) and Bargh himself later acknowledged "the claim that the activation of an attitude increases the accessibility of all similarly valenced representations in memory appears unlikely…" (Ferguson & Bargh, 2003, p.172).

The response competition mechanism

When describing processes that he believes might explain affective priming Fazio (2001) discusses the response competition mechanism, although it is sometimes known as the Stroop mechanism (Klauer & Musch, 2003). Fazio cites numerous researchers, including Klauer, RoBnagel, & Musch (1997), as nominating this mechanism.

The response competition mechanism postulates that the prime "readies" the response to the target (when the prime and target are the same) by preparing the evaluative pathway to the target. So if the prime and target are congruent (e.g. both positive words such as happy and wise) then the prime facilitates response to the target because the participant has already made one positive evaluation. If however, the prime and target are incongruent (e.g. happy and death) then the participant has to inhibit the positive evaluation pathway and activate the negative evaluation pathway to respond, which is why the response takes longer.

Fazio (2001) believed that the response competition mechanism might explain some of the affective priming effect. However he maintained that the results of some experiments that do not use target words with an obvious valence, such as Bargh et al. (1996a), might not be fully explained by the response competition mechanism. Bargh et al. (1996a) found the priming effect occurred when participants were asked to simply read the target word aloud rather than classify it as positive or negative; they also used target words that were not obviously positive or negative to avoid creating an evaluative mindset in their participants. Under these conditions evaluative pathways should not be activated and thus the priming effect would not be expected to occur. Fazio (2001) postulated that the spreading activation mechanism could explain why affective priming is still evident in these studies.

The affective-matching mechanism

Klauer and his colleagues (Klauer & Musch, 2003; Klauer et al., 1997; Klauer & Stern, 1992) described a third possible explanation of the affective priming effect, the affective-matching mechanism. Ferguson & Bargh (2003, p.173) describe this model as another example of a response competition explanation in which people attempt to integrate the prime and target word (if the prime is a noun and the target is an adjective) into a sentence (is object X adjective Y?). The affective-matching mechanism makes three assumptions:

1) The evaluations of both the prime and target are automatically activated and are compared for evaluative consistency (i.e. whether the valence of the prime and target match), regardless of the perceiver's current goals.
2) If two words are the same valence the perceiver feels a sense of plausibility, whereas evaluative inconsistency (where one word is positive and the other negative) leads to feelings of implausibility.
3) A spontaneously activated plausible feeling facilitates affirmative responses and inhibits negative responses, whereas an implausible feeling facilitates negative responses and inhibits affirmative responses.

Klauer & Musch (2003) used the example of the lexical decision task (participants decide if strings of letters were words or non-words) to explain the affective-matching mechanism. They said the model predicted that the affective priming effect (i.e. faster responses where the prime and target word are the same valence) would occur for word targets in a lexical decision task, even though the evaluations of prime and target are irrelevant to the goal, because words would require an affirmative response. Wentura (2000) did indeed show the pattern of responding predicted by the affective-matching mechanism. Furthermore, in another condition he reversed "yes" and "no" responses so that participants should respond "yes" to non-words and "no" to words. Under these circumstances the affective-matching theory predicts the priming effect should be reversed i.e. "no" responses to words should be faster when preceded by evaluatively inconsistent primes compared to consistent primes: i.e. if a participant was instructed to press a "yes" key if "disgusting" was a word then he or she should respond faster if they were primed with "cockroach" than if they were primed with "butterfly". However, if they were instructed to press a "no" key if "disgusting" was a word, then he or she should respond faster if they were primed with "butterfly" than if they were primed with "cockroach". This was indeed the case, supporting the idea that evaluative inconsistency facilitates negative responses via a feeling of implausibility whereas evaluative consistency inhibits negative responses.

However, while the affective-matching mechanism explains "yes"-"no" binary decisions, it cannot account for evaluative decisions (good/positive versus bad/negative decisions) because of the complexities regarding the confounding between the prime being consistent with the target and the prime being positive or negative i.e. any prime word can have two possible effects of a target word, one in terms of its valence (it could be positive or negative) and two, in terms of whether or not it matches the target (it could be consistent or inconsistent with the target) "it is... impossible to disentangle possible effects of prime valence from differences caused by affective matching" (Klauer & Musch, 2003, p.26).

This led Klauer & Musch (2003) to propose that "a large portion of the research (regarding mechanisms underlying automatic evaluation) is integrated by a model of the evaluative system that has at its core a process of preconscious evaluative activation" (p.27). They suggested at least two mechanisms, the response competition mechanism and the affective-matching mechanism, operating simultaneously and in parallel to explain the affective priming effect. They proposed that the response competition mechanism "explains most of the findings (but where it)... fails to account for affective priming in tasks without a strong evaluative component... the affective-matching model comes into play to explain affective priming in its wider contexts" (p.27).

4.2.3. Context and constructs of prime

Ferguson & Bargh (2003 p.174-5) noted that much of the work on automatic evaluations (which are measured by the APT) has assumed that they are the result of a "single affective tag" i.e. attitudes are solitary, previously stored, evaluative representations of an attitude object, that become activated when the attitude object is observed or thought about. They proposed that automatic evaluations are (i) constructive, i.e. they are formed of several different constructs or properties. They give the example of the category of dogs, pointing out people may have opposite associated attitudes towards the subcategories of "puppies" and "attack dogs", and; (ii) that automatic attitudes are context dependent. In this example an "attack dog" may be evaluated negatively if it seems likely to attack you, but positively if it is protecting you. Therefore the above explanations of what mechanisms underlie the affective priming effect may be incapable of giving a full explanation because they describe attitude objects as evoking only one emotional reaction and only one valence rating. In reality there may be more than one, and any explanation of affective priming needs to account for the fact that it is the salient construct of an attitude object and the context in which the attitude object is viewed that determines how it primes the target words.

A phenomenon known as list context effects may also explain some of the affective priming effect (Klauer & Musch, 2003). This refers to the extent to which an individual trial in the affective priming effect (i.e. evaluating one word as positive or negative, after being primed) is a local phenomenon, and to what extent it is dependent on the wider context i.e. where previous responses were "positive" to what degree will the current response will be facilitated if a positive response is required and inhibited if a negative response is required. Klauer & Musch (2003) reviewed work in this area and described the consistency proportion effect (i.e. the number of trials in an APT that require the same response). They summarized a series of unpublished experiments which are written in German, and also refer briefly to other studies e.g. Klauer et al. (1997). They concluded that consistency proportion moderates the affective priming effect but cannot override it. Additionally it appears to rely on learning processes that extend over a wider range of trials and require prime visibility to occur (i.e. the primes must not be masked and should be presented at speeds that the participant can see). Both the trial-by trial context (i.e. the trial preceding the current trial), and the global list context (i.e. the majority of the trials completed) in which the trials are embedded can modulate affective priming.

1.4.2.4. How automatically activated attitudes guide behaviour

Automatically activated attitudes have been shown to guide behaviour and perception of the environment in a number of studies (Ferguson & Bargh, 2004). For example Bargh, Chen, & Burrows (1996b) found participants who were waiting for the second stage of a task were more likely to interrupt a staged conversation between the experimenter and another person if they were primed for rudeness rather than politeness. The participants had been asked to unscramble sentences, and the unscrambled sentences contained words that primed either rudeness, politeness, or were neutral. In another experiment by Bargh et al. (1996b) participants who were semantically primed for old age (by unscrambling sentences) walked at a slower speed after completing the experiment than participants who were not so primed. Dijksterhuis & van Knippenberg (1998) asked participants to answer questions from the game Trivial Pursuit after priming participants with either a prime for intelligence (university professors) or non-intelligence (football hooligans). The participants primed for university professors performed better than control participants, and participants who were primed with football hooligans performed worse than control participants. Chartrand & Bargh (1999) found participants would unconsciously mimic the actions of a confederate (for example rubbing their face or shaking their foot when sitting with their legs crossed). Furthermore, in a second experiment participants reported preferring confederates who mimicked their (the participants) body posture and mannerisms to confederates who did not.

In addition to this Chartrand and Bargh (2000, cited as an unpublished manuscript in Bargh & Ferguson, 2000) found that priming participants with attitude objects that were positive or negative could effect their mood, with participants who were positively primed being in a significantly better mood to participants who were negatively primed.

Ferguson, Bargh, & Nayak (in press, cited in Bargh & Ferguson, 2000) conducted three experiments to show how automatically activated attitudes can effect social judgement. The first two experiments investigated how automatically activated attitudes can affect how people interpret stimuli. The first demonstrated that when participants are completing word-stems that can be completed to mean positive or negative words, priming them can effect if the word is completed to be positive or negative in meaning (e.g. "GREE_" could become "GREEN" or "GREED"). The second study found that priming positive or negative attitudes could influence whether participants defined homonyms such as "MEAN" in a positive or negative way. The third study explored whether or not automatic evaluations could effect how participants' interpret social behaviour. They primed participants with an attitude object and asked them to read ambiguous behaviour descriptions; participants tended to infer a trait that was consistent in valence with the attitude object that served as a prime. Bargh & Ferguson (2000) believed this demonstrated how one's initial automatic evaluation of a person (which could be based on their appearance) could effect how you interpret their behaviour and thus have a long lasting effect on your future interactions with that person. These studies clearly highlight the impact attitudes can have on our behaviour and perception of the environment, and thus emphasise the importance of automatically activated attitudes. If appearance can automatically shape attitudes there is clearly an opportunity for dental fluorosis to impact on an affected individual's interactions with other people.

1.4.2.5. The importance of affective priming and automatically activated attitudes

Attitudes can be automatically activated by priming, and this clearly has enormous potential to shape our perception and behaviour. By developing our understanding of when, why, and how attitudes are activated it may be possible to apply such knowledge to real life problems. For instance, in a recent experiment Arndt, Greenberg, & Cook (2002) raised the possibility that by measuring how certain worldview cognitions are primed when thinking about death it may be possible to determine which beliefs can be used in terror management i.e. in helping people to cope with their fears about death. They found that after priming participants to think about death (by asking them to list their thoughts about it) men were more likely to complete word stems with nationalistic words whereas women were more likely to use romantic words. They postulated that by measuring the accessibility of different constructs after mortality salience it may be possible to determine which of the beliefs people hold are used for terror management. These could then be used in experimental and clinical settings and "by applying therapies that encourage the application of particular belief and value structures, it may be possible to help individuals adopt more healthy and productive responses to such existential concerns" (p.322).

Walther (2002) found that attitude formation is not confined to the association between an attitude object and its evaluation; she found an attitude object can take on the evaluation of an attitude object it is associated with. The implications of this in a real-life scenario were highlighted by Castelli, Zogmaister, Smith, & Arcuri (2004) who looked at whether photographs of people serving as social exemplars of various categories would be automatically evaluated if participants had an existing opinion of the category, but there was no cue present to say that the photograph was a member of that category. They found evidence that it is possible to form an evaluation about a person based on an interaction with them, forget the interaction, yet for the evaluation to remain and still be activated upon seeing the person. The evaluation could then be misattributed to something else. They also found that our evaluations of people can guide our approach/avoidance behaviour.

In another study Fazio et al. (1995) used affective priming to develop the "*bona fide* pipeline", to show the extent to which individuals categorise others by race, and it has since been used in numerous studies, e.g. Fazio & Dunton (1997), Fazio & Hilden (2001), Olson & Fazio (2003), Towles-Schwen & Fazio

(2003), Olson & Fazio (2004b). Categorising people by race is just one form of automatic categorization; other studies have shown affective priming effects how objects are categorised. Smith, Fazio, & Cejka (1996) asked participants to respond to stimuli by pressing a "yes" or "no" key to answer questions regarding whether or not attitude objects belonged to certain categories. Some of the attitude objects could belong to more than one category (e.g. "yoghurt" could be categorised as a "health food" or a "dairy product"). They found that participants were faster to respond to categories that had been primed. So in the example of yogurt, whether we perceive it as a health food could influence our consumer decisions and therefore our diet.

A formal theory of how affective priming is important in how attitudes guide behaviour has been developed. The MODE model (Fazio, 1990a; Fazio & Towles-Schwen, 1999; Schuette & Fazio, 1995) (Fazio & Towles-Schwen, 1999; Schuette & Fazio, 1995) proposes that attitudes can be spontaneously activated and then guide behaviour in an automatic manner. The MODE model distinguishes between deliberative and spontaneous processes and proposes both can guide behaviour. Moreover, it suggests that opportunity and motivation to consciously consider one's attitudes determines which process becomes active and ultimately guides behaviour. That is, upon encountering an attitude object an individual's attitudes toward that object may be activated from memory, and influence the perception of that object and may even ultimately guide behaviour towards the object. However, with sufficient motivation and opportunity the individual can consider the object in terms of the attributes they know it to possess rather than relying on their previously stored summary evaluations of the object. Moreover, automatic and controlled processes can sometimes both be involved in judgements and behaviour. This was discussed in terms of racial prejudice, whereby an individual who is motivated to be nonprejudiced may (with sufficient opportunity) prompt efforts that mitigate a racially prejudiced automatically activated attitude and thereby avoid prejudiced behaviour.

The above studies highlight how important affective priming could be in influencing the lives of people, and support Fazio's (2001) belief that further investigation into the processes involved in affective priming, and the variables that moderate it, should prove valuable.

1.4.3. The Implicit Association Test (IAT)

The Implicit Association Test (IAT) was introduced by Greenwald et al. (1998). Like the APT (Fazio et al. 1986) the IAT utilises response latency to implicitly measure attitudes. But where the APT uses an attitude object to prime a valence judgement the IAT measures association between constructs by forcing participants to sort items belonging to four constructs using two responses, so that two constructs have to share a response.

A typical IAT would comprise five stages (see Table 1.1). The first stage involves target-concept discrimination. In this stage participants are shown words on a monitor which, for instance, would be either insects or flowers (e.g. ant, rose). The words are shown one at a time and participants categorise them as insects or flowers by pressing one of two response keys, for instance flowers might be assigned to key 1, and insects to key 2. The second stage is the associated attribute discrimination, in which participants are shown words taken from two new categories, such as pleasant and unpleasant (e.g. lucky, poison). As before the words are displayed one at a time, and as before participants categorise them using one of two response keys, for instance, assigning unpleasant items to key 1 and pleasant items to key 2. The third stage is the initial combined task. In this stage participants are asked to press key 1 when shown a word that is a flower or is unpleasant, and asked to press key 2 when shown a word that is an insect or is pleasant. The fourth stage is the reversed target-concept discrimination stage. In this stage participants get used to assigning the opposite keys to two of the concepts, for example flowers would now be assigned to key 2, and insects assigned to key 1. The fifth stage is the reversed combined task. In this stage participants are asked to repeat stage three but this time to use the reversed keys they practised in stage four. So they would press key 1 when shown a word that is an insect or which is unpleasant, and key 2 when shown a word that is a flower or is pleasant. The response latency of stage five is then subtracted from the response latency of stage three to give the IAT score.

The IAT is expected to reveal associations between the concepts by showing faster response latencies for combinations that are evaluatively similar. In the present example faster responses would be expected for flowers+pleasant and insects+unpleasant than for insects+pleasant and flowers+unpleasant, because flowers are associated with pleasant words and insects with unpleasant words.

Table 1.1. Example of an IAT

Stage	Concepts for key 1	Concepts for key 2
1	Flower	Insect
2	Unpleasant	Pleasant
3	Flower + Unpleasant	Insect + Pleasant
4	Insect	Flower
5	Insect + Unpleasant	Flower + Pleasant

Table adapted from Greenwald & Nosek (2001).

1.4.3.1. The generality of the IAT

The IAT has been widely used in variety of studies, of which the most common have been examinations of prejudice, and in particular racial prejudice. Indeed, in the first IAT paper published, Greenwald et al. (1998) used the IAT to identify prejudice among Korean and Japanese participants, and among white participants who revealed a preference for typically white names over typically black names.

Subsequently, Phelps et al. (2000) showed that amygdala activation (the amygdala is a part of the brain used in emotional learning) in white participants was higher when shown unfamiliar black faces than when shown unfamiliar white faces (which indicates more fear, discomfort etc). When attitudes were measured using a paper based explicit measure (the Modern Racism Scale, McConahay, 1986; McConahay, Hardee, & Batts, 1981) participants did not appear racist; however the IAT showed participants had a preference for white faces over black faces. The IAT scores correlated with the strength of the amygdala activation for Black-versus-White faces, but scores on the Modern Racism Scale did not.

Other studies have implicitly revealed religious (Christian and Jewish), age (young versus old), and national (American versus Soviet) prejudice (Rudman et al., 1999), prejudice between East and West-Germans (Kuehnen, Schiessl, Bauer, Paulig, Poehlmann, & Schmidthals, 2001), prejudice against homosexuals (Banse, Seise, & Zerbes, 2001), and gender stereotypes for mathematics (Nosek, Banaji, & Greenwald, 2002)

However, not all of the work on the IAT has been used to study prejudice. Greenwald & Farnham (2000) conducted three experiments that compared IATs designed to measure self-esteem and self-concept with explicit measures. Their findings led them to conclude that the IAT is stable enough to be used in research settings and that it taps different constructs from self-report measures. This is discussed in more detail in Section 1.4.3.3.

de Jong, Pasman, Kindt, & van den Hout (2001) attempted to use the IAT to distinguish between high and low socially anxious women (participants were identified as high or low in social anxiety prior to selection in the study). Their IAT used the constructs of social situations versus neutral words (e.g. date and hall), and positive versus negative outcomes (e.g. compliment versus rejection), and they predicted socially anxious women would show a greater implicit association between social situations and negative outcomes. However, their results did not show any statistically significant differences between the women. Although in a later study de Jong (2002) successfully used the IAT to investigate if socially anxious people had lower self-esteem than non-socially anxious people, and if the esteem they held others in was different from non-socially anxious people. He found that socially anxious individuals were characterised by a relatively positive image of others and it may be this reduced tendency of self-favouring that is pivotal to social anxiety. de Jong (2002) proposed these findings casts doubt on the idea that negative self-evaluations *per se* are an important feature of social anxiety.

In addition to this the IAT has been used, with varying degrees of success, to measure attitudes to shyness (Asendorpf et al., 2002), smoking (Swanson, Rudman, & Greenwald, 2001), condom use (Marsh et al., 2001), high fat foods (Roefs & Jansen, 2002), and to distinguish psychopaths who have committed murder from non-psychopathic murders and other offenders (Gray, MacCulloch, Smith, Morris, & Snowden, 2003).

1.4.3.2. What does the IAT measure?

Since its introduction, one of the most commonly voiced concerns over the IAT has been that rather than measuring the respondents' personal evaluations it measures their knowledge of cultural associations between the constructs in the IAT. In fact it was a concern that Greenwald et al. (1998) noted themselves when they introduced the IAT (although they also pointed out that familiarity couldn't account for all of their findings). For example in an

IAT that measures racial prejudice, participants may be familiar with cultural stereotypes that associate black people with negative descriptions, and therefore may be quicker at mapping black names or faces to the same key as negative descriptions than to the same key as positive descriptions, even though they do not hold nor endorse such attitudes themselves.

A number of studies have demonstrated how cultural knowledge and familiarity influence IAT scores. Brendl, Markman, & Messner (2001) ran three IATs comparing non-words (unfamiliar stimulus) to positive and negative categories. They found that non-words were more negatively evaluated than both positive and negative words even if participants were told the non-words were foreign words that were positive in meaning. It could be that unfamiliar items such as non-words have a pre-stored negative valence or that real words appear to be preferred due to being more familiar (e.g. the mere exposure effect, Zajonc, 1968). Karpinski & Hilton (2001) found the IAT appeared to measure participants' cultural knowledge about apples and candy bars, rather than their personal evaluations. They asked participants to do an IAT (and equivalent explicit measures) to determine if they preferred apples or candy bars. Participants were then given the choice of an apple or candy bar. The explicit attitude measures predicted participants' choice, but the IAT did not. In another study Karpinski & Hilton (2001) manipulated associations between experimental concepts to show that environmental, rather than internal, associations determine IAT score. Participants completed a youth-elderly IAT, and then in an apparently unrelated task completed a memory test that either reinforced the youth = positive, elderly = negative stereotype, or contradicted it. After this manipulation participants repeated the youth-elderly IAT. Results showed that the participants in the elderly = good condition, decreased in their preference for youth after exposure to the memory test according to the IAT. Karpinski & Hilton (2001) took this as support for an "environmental association interpretation of the IAT" (p.774).

However, while there is clearly evidence that familiarity could explain some of the IAT effect there are a number of studies that have controlled for the effects of familiarity and still found the IAT effect. Rudman et al. (1999), for example looked at national prejudice using the surnames of Soviet and American leaders. They controlled for familiarity by using familiar and unfamiliar leaders from both countries, and combining them to produce four IATs: (1) familiar U.S. leaders versus familiar Soviet leaders, (2) unfamiliar U.S. leaders versus unfamiliar Soviet leaders, (3) familiar U.S. leaders versus unfamiliar Soviet leaders, (4) unfamiliar U.S. leaders versus familiar Soviet leaders, (participants only completed two of the four IATs each). All four IATs showed (the American) participants preferred U.S. leaders, and IATs 3 and 4 (where you would expect the differences caused by familiarity to be most evident) had almost identical effect sizes. In short their findings showed implicit preferences for participants own country regardless of their familiarity with the stimulus.

Banse et al.'s (2001) finding that heterosexuals showed more negative attitudes to homosexuals than did homosexuals suggests that the IAT measures the opinions of the individual completing it rather than their knowledge of a cultural stereotype; otherwise you would expect no differences in the IATs of heterosexuals and homosexuals. Dasgupta, McGhee, Greenwald, & Banaji (2001) used two IATs designed to avoid a finding caused by familiarity, the two provisions they used avoid familiarity were; (i) a procedure that measured and statistically controlled for participants familiarity with the white and black names used in the IAT test of racial prejudice, and (ii) photographs of Black and White faces instead of names, so that participants were equally unfamiliar with all of the faces. Yet the IAT still showed the expected results.

Perhaps the debate surrounding the effects of familiarity is about to resolved. In a recent paper Olson & Fazio (2004a) introduced a "personalised IAT" which is intended to focus the IAT more on the evaluations of the respondent by removing cultural knowledge from their responses. They conducted a series of experiments to test the personalized IAT; in the first participants undertook a traditional IAT and a personalized IAT. The personalized IAT had three differences to the traditional IAT:

1) Instead of using the category labels "pleasant" and "unpleasant" when participants categorized the evaluative words, they used the labels "I like" and "I don't like".
2) The evaluative items they used were no longer normative as in the traditional IAT, but more ambiguous. That is instead of using words that almost everybody categorizes as pleasant or unpleasant, (e.g. love, sickness), they used words that people have varying opinions about (e.g. coffee, country music).

3) If participants made a mistake on the personalized IAT they weren't given error feedback, as they were in the traditional IAT.

They predicted that the personalized IAT would identify less racial prejudice than the traditional IAT; their reasoning was that the traditional IAT was identifying non-prejudiced participants as prejudiced because of their cultural knowledge of the negative stereotypes about black people, but the personalized IAT should reveal the participants own attitudes. As expected the personalized IAT did reveal less prejudice. In three further studies Olson & Fazio (2004a) experimented with the criteria required for a personalised IAT and suggested that it is unnecessary to change the evaluative items for less normative words, and that changing the evaluative category labels and removing error feedback is sufficient. They also found further support for the beneficial effects of the personalised IAT in the domains of predicting apple versus candy bar preference and political attitudes.

The debate over whether the IAT measures personal preference or familiarity is not straightforward, since mere exposure effect (Zajonc, 1968) postulates that we prefer things we are familiar with to those we are unfamiliar with. If this is the case then we should expect some confounding between personal preference and familiarity (although it is possible to be familiar with things we don't like). The personalized IAT may settle this issue, but until further research is done with the personalized IAT it is not possible to say to what extent the effects of familiarity are a problem for the IAT.

Although the effects of familiarity on the IAT effect is the most widely researched and discussed issue, there have been a number of other concerns about the extent to which the IAT measures attitudes. In particular the effects of task-switching, method-specific variance, and ground-figure asymmetry have been shown to influence IAT scores. There have also been concerns about whether the IAT measures cognitive or emotional processes, the effects of item confounding and the stability of the IAT. These issues are now discussed.

Task-switching

Mierke & Klauer (2001) found that the IAT effect can be at least partly explained by task-switching (for example, in one trial you judge if a word is positive or negative in meaning, but in the next trial you judge if a word is a flower or an insect), that is, some of the time difference between blocks three and five is due to task switching rather than to associations held by the participants. Their first experiment was a typical flower/insect IAT and they found the usual results, (i.e. flowers were preferred to insects) but they analysed beyond that and also found this was caused by a greater cost of task switching in the incompatible block (insects+pleasant and flowers+unpleasant) compared to the compatible block (flowers+pleasant and insects+unpleasant). They predicted that if participants pressed the same button for two consecutive trials there would be no time cost in the compatible block (e.g. if they pressed they left button twice consecutively, once for a positive word and once for a flower), but that there would be a time cost in the incompatible block (e.g. if they pressed the left button twice consecutively, once for a positive word and once for an insect). However, this was not the case; there was a time cost even in the compatible block (the time cost is in comparison to pressing, for example, the left button twice in consecutive trials, both times because the word is positive in meaning).

Experiment 2 found that cueing participants to which button they should press in the upcoming trial reduced task-switch effects. They cued participants by putting a symbol to the left of the word if participants should press the left button, and to the right of the word if participants should press the right button. The results showed that cueing participants reduced the cost of task switching, and this reduced the size of the IAT effect, which means that at least some of the IAT effect can be explained by the time costs of task switching.

Method specific variance

In a later study Mierke & Klauer (2003) found the IAT contains method specific variance which can explain some of its high test-retest reliability and internal consistency. They discovered that the IAT effect can be found in the absence of a pre-existing association between the response categories. In other words, the IAT is believed to work, because, for example, flowers have more positive associations than insects, and insects have more negative associations than flowers, but if there are no associations in existence then the IAT effect should not be found. However, Mierke & Klauer (2003) ran a study in which the IAT used sizes and colours of shapes as its constructs. There is no reason why a size (big or small) should be more closely associated with a colour (red or blue), but they imposed an artificial relationship between blue and large and red and small. The IAT effect was found even though there was no pre-existing association, i.e. there was method specific variance.

In another experiment participants completed two IATs one after the after, one of which was the flower-insect IAT and one of which was the colours-sizes IAT. Then they looked for correlations in these two IATs within individuals, correlation would demonstrate that the method specific variance evident in the colours-sizes IAT was also present in the flower-insect IAT since the underlying constructs in the two IATs were completely unrelated and could only be explained by inter-individual differences. The two IAT scores did correlate, so it therefore seems that variance found by the IAT can be explained in part by method specific variance and cannot be entirely credited to the associations between the objects/evaluations being judged. However, Mierke & Klauer (2003) found that the method specific variance could be reduced using a scoring algorithm proposed by Greenwald, Nosek, & Banaji (2003).

Ground-figure asymmetry

As described above the basic idea of the IAT is that the IAT effect is caused by associations e.g. there is a stronger association between flowers and pleasant than insects and pleasant, and there is a stronger association between insects and unpleasant than between flowers and unpleasant. Rothermund & Wentura (2001) found that IAT effects could be partly caused by something other than associations, namely figure ground asymmetries. This means that when participants are doing the IAT instead of focusing on each of two categories equally, they focus their attention on deciding if each word does or does not belong to one of the categories and class anything else as being in the other category i.e. they do a visual search task rather than a binary classification task. Participants are more likely to focus on the category that is more unfamiliar or negative of the two since it is more noticeable, i.e. it "pops out" at them. In their analogy the more noticeable category is termed "figure" and the less noticeable category termed "ground". So, for example, instead of focusing on flowers and insects equally they focus on insects, then make a visual search for insects and just use the alternative response for anything that isn't an insect, rather than making an "insect or flower?" judgement. This means that when it come to the mixed response section of the IAT (e.g., insect or flower and pleasant or unpleasant) they might respond in a way that associates insects with unpleasant and flowers with pleasant, but the only thing insects have in common with unpleasant is that they are the designated "figures" in the visual search and are not associated with each other in a meaningful way, and ditto for the two "grounds" flowers and pleasant.

Does the IAT measure cognitive or emotional processes?

Plessner & Banse (2001) raised the concern that the cognitive and/or affective mechanisms underlying the IAT had received little attention. Since then Cunningham, Johnson, Gatenby, Gore, & Banaji (2003), like Phelps et al. (2000), have used functional Magnetic Resonance Imaging to provide evidence that the IAT is measuring affective rather than cognitive processes. Cunningham et al. (2003) studied which areas of the brain were active during different tasks. Participants were shown the names of famous people and made either evaluative judgements, or non-evaluative judgements about them. Their findings suggested that there are separate brain systems involved in automatic or perceptual processing and in controlled or reflective processing. Brain areas associated with automatic processing were sensitive to simple valence information, whereas brain areas associated with controlled processing were sensitive to the complexity of the information being processed, such as ambivalent evaluations, rather than to its valence.

Item confounding

When completing an IAT participants classify items under different constructs (e.g. they classify "ant" under the construct of insect), and the IAT then measures the extent to which they associate certain constructs with other each other based on their response latency. However, the items used in the IAT may effect how closely associated the constructs are found to be. De Houwer (2001) used the concepts of British names/foreigner names and positive/negative to look at whether the IAT measured the difference between the valences of the target concepts e.g. "British" and "Foreigner", which he termed the relevant feature account of the IAT. Or whether it measured the mean difference in the valences of the actual target items e.g. "Margaret Thatcher" and "Albert Einstein", which he termed the irrelevant feature account. His results supported the relevant feature account, suggesting that participants simply classify the target items to the predefined categories without the valence of the individual items affecting their response latency.

However, in a similar study, Steffens & Plewe (2001) examined how the items used can have a confounding influence on the IAT effect. They used the concepts male/female and positive/negative. They expected to reveal an implicit preference for females, as past work had found this. Participants completed the IAT under one of two experimental conditions, either the positive words were stereotypically female attributes and the negative words stereotypically male attributes (e.g. beautiful, violent), or, in the second condition, the positive words were stereotypically male attributes and the negative words stereotypically female attributes (e.g. gentleman-like and bitchy). Both conditions showed the expected IAT effect (female preferred to male), but it was much bigger in the condition where the positive words were stereotypically female attributes than in the condition where the positive words were stereotypically male attributes (although not quite significantly bigger $p<.07$). It would therefore appear the items used to represent the constructs being measured can affect the size of the IAT effect. The finding of De Houwer (2001) supporting the relevant feature account, would still appear to be pertinent, given that the IAT was still in the expected direction in the Steffens & Plewe (2001) study, but it is clear that care needs to be taken when using the IAT as its effect size can be moderated by the items used to represent the target concepts.

Stability

Kuehnen et al. (2001) found that people's attitudes, even when measured implicitly may not be "robust" (i.e. constant), but rather can change according to the context or situation. This variability has implications for how robust the IAT is. They conclude that "the robustness of the IAT is limited by the robustness of attitudes" (p.143).

de Jong (2002) raised concerns about the stability of the IAT. He used the IAT to investigate if socially anxious people have lower self-esteem, and if socially anxious people have different views about the esteem they hold others in. Their procedure was completed twice, three months apart. They showed that while the main pattern of the IAT was stable across time, the absolute IAT indices were highly variable across the two assessments.

1.4.3.3. The IAT and explicit measures

Many of the studies that have investigated the IAT have also used equivalent explicit measures of attitude to allow a comparison with the IAT. There have been a range of findings in terms of how well the IAT correlates with explicit measures of attitude, but it is widely believed that the IAT often (but not always) measures variance not accounted for with explicit measures (Greenwald & Nosek, 2001). The explicit measures that are most commonly used when studying the IAT are semantic differential scales and feeling thermometers (see Section 1.4.1). The results given by this type of scale can be combined to create an overall score which can then be compared to the IAT score. Other explicit measures have been used, including questionnaires utilising Likert or semantic differential scales which also provide a single overall score for comparison with the IAT e.g. Greenwald et al. (1998) asked participants to complete the Modern Racism Scale. Across three experiments that measured national and racial prejudice in addition to attitudes towards flowers and insects, Greenwald et al. (1998) concluded that explicit measures (semantic differential scales and feeling thermometers) correlated with each other better than they correlated with the IATs, and that their results should be taken as "evidence for divergence of the constructs represented by implicit versus explicit attitude measures" (p.1477).

One of the advantages of implicit measures over self-reported measures is that they are less open to self-presentation strategies. This may be particularly important where the attitudes being measured are socially contentious. Banse et al. (2001), for example, examined how easy it was for participants for fake their responses on explicit measures and the IAT. They developed a homosexuality IAT designed to measure prejudice, and also developed explicit measures of attitudes towards homosexuality and motivation to control prejudiced behaviour. They then manipulated how motivated participants were to fake positive attitudes to homosexuals. There were four conditions, (i) no request for participants to modify their attitude, (ii) the demand condition which involved a small piece of text that implied it was wrong to discriminate against homosexuals, (iii) the persuasion condition which involved the same text as condition (ii), plus some extra text encouraging positive attitudes to homosexuals and arguing against discriminating against them, (iv) participants were asked to deliberately fake positive attitudes to homosexuals. The results showed it was possible for participants to control explicitly measured attitudes and the stronger the message encouraging positive attitudes to homosexuals the more positive attitudes the participants reported. However, it was not possible for participants to control the responses in the IAT, and there were no significant differences between the four groups of IAT scores.

de Jong (2002) noted that because implicit attitude measures are better at avoiding self-presentation strategies it was beneficial to supplement self-report measures with indirect measures of self-evaluation, in a clinical setting. Adopting this approach he found high-anxiety participants had lower self-esteem than low-anxiety participants, if self-esteem was explicitly measured, but when self-esteem was measured implicitly on the IAT they did not show lower self-esteem. This led him to conclude that self-reported low self-esteem in high-anxiety individuals "more accurately reflects the strength of their self-presentation concerns than their self-esteem per se" (p.507). Although in an earlier study de Jong et al. (2001) did find the IAT was not as effective as established self-report measures at distinguishing between high and low socially anxious women. This suggests the IAT is perhaps suitable for use in clinical settings only if it is used in conjunction with established explicit measures, but not as a replacement for them.

A number of studies have found evidence that although the IAT measures variance not accounted for by explicit measures, it does still correlate with explicit measures, e.g. Cunningham et al. (2001) who compared measures of racial prejudice. In another example, Greenwald & Farnham (2000), compared how well explicit measures and the IAT measured self-esteem and self-concept. In their first experiment, which investigated self-esteem, participants completed a number of explicit measures and two IATs. The IATs correlated with each other significantly and the explicit measures correlated with each other significantly, but the IAT and the explicit measures only correlated weakly. In a second experiment they used two IATs to investigate gender self-concept and participants also completed explicit measures of gender of self-concept. The IATs were three times more effective at identifying gender differences than the explicit measures, and again there were weak but positive correlations between the implicit and explicit measures. A third study showed that the IAT was more successful than explicit measures at identifying predicted interactions between self-esteem and success or failure. They concluded that the IAT; (a) has stability and validity to justify its use in research settings, and (b) defines constructs different from, although correlated with, constructs that nominally measure the same thing using explicit measures.

Swanson et al. (2001) found evidence that attitudes towards stigmatised behaviours, such as smoking, can have a moderating effect on the dissonance between implicit and explicit measures. They used the IAT and explicit measures to measure attitudes to smoking and found that smokers rated smoking preferentially when measured explicitly but not when their attitude was measured implicitly (by the IAT), while non-smokers rated smoking negatively both explicitly and implicitly. Experiment 2 was designed to show implicit and explicit attitudes correlate more strongly for behaviours that are not stigmatised (vegetarianism) than for stigmatised behaviours (smoking). The IATs and the explicit measures all showed that vegetarians preferred, and identified themselves more, with other protein compared to meat, but omnivores preferred and identified more with meat. Furthermore, another IAT showed that both smokers and non-smokers held equally favourable implicit attitudes to smoking compared to stealing. A separate IAT showed that smokers and non-smokers identified more with smoking than stealing but smokers identified significantly more strongly with it than did non-smokers. Explicit measures also found that smokers preferred smoking significantly more to stealing than did non-smokers. Therefore it does appear that there is a greater correlation between implicit and explicit measures of attitude when the behaviour being rated is not stigmatised, compared to stigmatised behaviours. Experiment 2 also suggests the gap between smokers' attitudes to smoking and their behaviour is not caused by them not identifying themselves with smoking.

There are, however, studies which have found only very weak correlations between explicit measures and the IAT. Rudman et al. (1999) compared their Jewish/Christian IAT with explicit measures. They found that "the IAT's relationships with subject religion and the acquaintance measure were robust, whereas the relationships shown between the IAT and direct measures of pro-Semitism were relatively modest" (p.448-9). In addition to this, Rudman et al.'s (1999) investigation of ageism and age stereotyping, asked participants to complete two IATs as well as explicit measures. The results showed that relationships between explicit and implicit measures were generally weak. Phelps et al. (2000) found that in a study of racial prejudice the IAT correlated with other implicit measures (amygdala activation and potentiated startle response), but an explicit measure (the Modern Racism Scale) did not.

There have been other studies of the how well the IAT and explicit measures predict behaviour. Asendorpf et al. (2002) developed an IAT of participants' self-concept of shyness. In addition to the IAT participants gave self-ratings of their shyness and other personality traits and were filmed talking to a confederate in a situation likely to induce shyness (they were speaking to an attractive member of the opposite sex, knew they were being filmed, and knew they would be asked to rate each others personality). Their behaviour was then analysed by independent judges in order to identify the duration of spontaneous shy behaviour (e.g. their body tension and facial and body adaptors such as self-stimulation of the face/body) and controlled shy behaviour (e.g. the speech of the participant, and movements or gestures). They found that the IAT could be used to detect differences between individuals in their self-concept of shyness, and that the IAT appeared to be separate from their explicit (self-rated) self-concept of shyness. The IAT predicted spontaneous shy behaviour but not controlled shy behaviour and the explicit measure predicted controlled shy behaviour but not but spontaneous shy-behaviour. The implicit and explicit measures correlated moderately.

Marsh et al. (2001) looked at how several explicit and implicit measures of attitudes towards condom use predicted reported condom use. They expected that when having sex with a steady partner there is an element of deliberate intention and therefore explicit measures should predict condom use, as intentions, like self-reported attitudes are at the conscious level. Conversely, they predicted having sex with a casual partner would be more spontaneous, and therefore participants would follow their affective feelings to condoms rather than their deliberate intentions, in which case implicit attitudes would be a more reliable prediction of behaviour. They found explicit measures tended to predict condom use with a steady partner, but not with a casual partner. By contrast, condom use with casual partners, was only associated with one of three implicit measures, and this was a relatively weak association. An IAT designed to measure affect to condom use did correlate with reported condom use with casual partners. However, an IAT designed to measure self-identification with condom use, and an APT (which primed positive/negative images using condom/non-condom images) did not predict reported condom use with casual partners. The two IATs had minimal correlation with each other, and neither correlated with the APT. None of the implicit measures were expected to, nor did predict condom use with a steady partner.

Clearly then there is evidence that while explicit measures of attitude predict controlled deliberative behaviour, the IAT may be more successful at predicting spontaneous behaviour.

1.4.4. The current research

The primary aim of the research conducted in this thesis is to examine how social judgements made about individuals vary according to the level of fluorosis. Therefore the question is which of the two implicit measures is better suited to the current goal, or, if neither is ideal, would a different task, adopting elements from one or both of them be optimal? To make this decision the literature on both the APT and IAT has been considered and there does not appear to be a precedent in either measure of examining social judgements based on dental appearance. These measures appear to have been used in very few studies of social judgements based on appearance, beyond the racial prejudice studies. Therefore it was not possible to simply choose an existing methodology and adapt it to the current aim, and instead thought needed to be given to the design of an implicit measure of attitudes to dental fluorosis.

The design that was decided on is described in detail in Chapter 3, and the rationale behind this choice is discussed in Chapter 5. This chapter has sought to describe, using relevant literature, the social importance of appearance, and to highlight how dental fluorosis might negatively impact on social judgements made about an individual affected with this condition. It then discussed possible methods that could be used to measure the impact of dental fluorosis on such social judgements. Having established that there is justification for developing an implicit measure of attitudes to fluorosis, the following chapters describe experiments in which firstly, the materials needed to create an implicit measure were developed, and then two studies in which this methodology was applied. Finally the implications of these studies are discussed.

1.5. The aims of the current research

The aims of the current research were to:

(1) Identify characteristics attributed to others, in relation to severity of dental fluorosis.

(2) Measure the nature and attitude strength of attributions made about others with differing levels of fluorosis using both explicit and implicit methodologies.

(3) Measure the nature and attitude strength of attributions made about others with dental caries using both explicit and implicit methodologies.

(4) Compare attitudes to fluorosis and dental caries.

Chapter 2. The Public's Description of Dental Fluorosis

2.1. Introduction

This chapter describes the first study undertaken in the course of this thesis. It begins by explaining how the standardised images used in all subsequent studies were produced and how the descriptions used in this first study were selected. It then reports how the images and descriptions were used to assess the general public's perception of fluorosis, when viewing extra- and intra-oral photographs. The descriptions identified as relevant to fluorosis were used in later studies.

2.2. Aims and objectives

Having generated standardised images of fluorosis that ranged in both the degree of severity (i.e. none, mild, moderate, and severe) and view (i.e. extra-oral and intra-oral) the objectives of the study were to identify how participants:

1) Described the images in their own words.
2) Selected characteristics from a pre-defined list to describe the images.
3) Varied in their descriptions and selections according to whether they viewed the extra-oral or intra-oral images and according to whether or not they received cueing to pay particular attention to the mouth when viewing the extra-oral images.
4) Having identified which descriptions varied across levels of fluorosis, to use these in subsequent studies.

2.3. Method, materials and procedure

2.3.1. Production of standardised images

The initial stage of the investigation involved the production of standardised images. Extra and intra-oral photographs were taken of student volunteers using a Fujifilm Finepix S1 Pro camera (Fujifilm, Tokyo), a Nikon Macro Speedlight SB-29, and an AF Micro Nikkor 105mm lens (Nikon, Tokyo).

Volunteers were photographed from two perspectives: A frontal view smiling with the teeth showing, and a standard intra-oral view of the anterior teeth (conventional lip retractors were used). These images were exposed under standardised lighting conditions to ensure the colour of the teeth was consistent across the two views.

The volunteers who posed for these photographs were dental students of various years at the University of Wales College of Medicine. They were approached in either the student common room, in their tutorials, or at the end of their lectures and asked if they would pose for the photographs. Consent for use and manipulation of the images was obtained and a small fee was paid to reimburse the volunteers for their participation.

The models selected for these photographs were not of unusual appearance, with healthy teeth that did not show any dental abnormalities. Approximately 20 students posed for photographs to provide a pool of images for digital manipulation. Four of these images (one male, one female from both the extra- and intra-oral view) were subsequently modified using Adobe

PhotoShop version 6, to simulate normal enamel, mild, moderate, and severe fluorosis based on Dean's Index (Dean, 1934). These photographs were produced as life size (extra-oral full face showing the teeth) and 4 x 6-inch (intra-oral) colour photographs. The photographs are shown in Appendix B.

As a result, standardised extra- and intra-oral photographs (of the same individuals) displaying a range of severity of dental fluorosis were available for assessment as described in Section 2.3.4.

2.3.2. Production of a list of characteristics

To determine participants' assessment of characteristics associated with the standardised images, a tick-list was developed by:

(1) Selecting 200 personality trait words from a study by Anderson (1968). He asked participants to rate 555 personality traits for meaningfulness and valence (i.e. how positive or negative they were in meaning). He then identified 200 words that were identified as being most meaningful to people, these words were selected for use in this study.

(2) These 200 words were reduced to 90, using systematic selection to take a sample of the full range of valence. That is, Anderson (1968) listed the words in order of valence (i.e. from the word rated as being most positive to the word rated as being most negative), therefore words were selected at equal intervals to cover the full list and thereby the full range of valence.

(3) The words generated by Anderson (1968) were personality traits, and since previous work on perceptions of dental appearance included characteristics related to aesthetics (Eli et al., 2001; Shaw, 1981) ten

further aesthetic-related characteristics were added to the list. The experimenter selected these from the work by Shaw (1981) and Eli et al. (2001).

(4) The resulting 100 word list was reduced to 50 by removing words the experimenter considered to be least meaningful or repetitive of other words.

(5) The order of the words in the list was randomised using Excel (Microsoft; Redmond, WA) and SPSS (SPSS Inc; Chicago, Illinois) software.

This list is presented in Appendices C 1 - C 3.

The following section describes how the images and the tick list were used to examine perceptions of fluorosis.

2.3.3. Recruitment of participants

Participants attending a Sports and Leisure Centre were approached and asked to volunteer for the study. They were approached either in the main foyer, in the restaurant and bar, or as they passed through the main corridor. These participants provided a broader range of ages and backgrounds than would have been possible had the sample been drawn solely from a student population. Those who agreed to participate were given a verbal description of the experimental procedure, an information sheet containing instructions and a leaflet describing the background of the study (Appendices C 4 - C 7). Informed consent was obtained (Appendix C 8) and participants completed a demographic information sheet (Appendix C 9).

2.3.4. Description of photographs

Participants viewed the photographs in two stages. In Stage 1, participants viewed four photographs, one at a time. They viewed the same individual (either the male or the female) from the same perspective (extra-oral or intra-oral) with normal enamel, and with mild, moderate and severe fluorosis. They wrote their description of each photograph on a separate piece of paper, describing each image in their own words, and after each description they were asked to indicate, using a "+" or "-" sign, whether they regarded that word as being positive or negative in meaning.

Next, in Stage 2 of the study, they were asked to view the same photographs for a second time and on this occasion endorse adjectives from the 50-item descriptor list (Section 2.3.2) that they felt described the individual in each photograph.

The details of the experimental protocol followed by participants are shown in Figure 2.1 and were as follows. Firstly, participants were systematically allocated to one of three experimental conditions. One group viewed extra-oral photographs only; the second viewed the same photographs but were asked to pay particular attention to the mouth, whilst the final group viewed only the intra-oral images. Groups were constructed so that they contained equal numbers of males and females.

Participants viewed the photographs in order of increasing fluorosis severity, although the point in the sequence at which they began viewing was systematically varied; that is approximately a quarter viewed the photograph with no fluorosis first, a quarter began with mild fluorosis, and so on. Participants were allocated to ensure that 45 viewed male images and 45 female images.

Figure 2.1. The study procedure of the first study

Condition 1. 30 participants viewed extra-oral photographs.	**Condition 2.** 30 participants viewed extra-oral photographs and were asked to pay particular attention to the mouth.	**Condition 3.** 30 participants viewed intra-oral photographs
↓	↓	↓

Stage 1

Participants' viewed either the photographs of the male or the female, one at a time. All participants' viewed the photographs in the sequence: normal teeth → mild → moderate → severe fluorosis; however where the sequence began was systematically varied across participants.

↓

Participants wrote down their own spontaneous descriptions of the characteristics of the person in each photograph. They indicated whether they regarded each description as being positive or negative in meaning.

↓

Stage 2

Participants viewed each photograph again in the same order as in Stage 1.

↓

Participants endorsed descriptions they believed were relevant to each photograph, from a list of 50 characteristics.

This protocol resulted in two data sets for analysis; the characteristics spontaneously generated by the participants themselves, and the characteristics they endorsed from the tick list.

2.3.5. Pilot study

This procedure was reviewed after it was administered to the first 10 participants. Based on the data generated it was decided that the design of the study would be sufficiently effective and therefore no alterations were made to the study and the 10 participants were included in the main analysis.

2.3.6. Ethical approval

The study was approved by the Bro-Taf local research and ethics committee on the 23rd of July 2002.

2.3.7. Analysis

The large number of descriptions spontaneously generated in Stage 1 of the study precluded individual analysis. It was therefore decided to collapse words of a similar meaning into groups and to analyse the number of endorsements each group received at each level of fluorosis. Each group was assigned a titular characteristic that best represented the meaning of the words in that group. After the descriptions were grouped, a frequency count was taken of the number of times a word from each group was used, at each level of fluorosis. If more than one description from the same group was used by the same participant, at the same level of fluorosis, only one spontaneous use/endorsement of that group was counted. Comparisons between these frequencies were performed on groups that had at least 20 endorsements, using the Chi-square statistic.

The descriptions on the tick list (in Stage 2) were also reduced and analysed in this way, but were analysed separately from the spontaneously generated characteristics.

2.4. Results

2.4.1. Participants

The study sought the views of a convenience sample of 90 volunteers. These individuals were recruited from a Sports and Leisure Centre (Section 2.3.3). Equal numbers of males and females (45 of each) participated; they ranged in age from 18-66 years old with a mean age of 31.6 ($\pm$13.0) years old.

2.4.2. Data preparation

Prior to presenting the results of the study it is necessary to describe further the steps taken to prepare the data for analysis. The preparation is summarised in Figure 2.2.

Figure 2.2. Preparation of data for analysis in the first study

Data from Stage 1.	Data from Stage 2.
170 descriptions were generated by 90 participants.	The tick list was endorsed by 90 participants.
The 170 descriptions were collapsed into groups, based on how similar they were in meaning. This was done by eight participants who did not take part in the assessment of the photographs.	The 50 descriptions on the tick list were collapsed into groups, based on how similar they were in meaning. This was done by eight participants who did not take part in the assessment of the photographs.
The eight participants then gave a title to each group that best represented what the **descriptions** in that group meant. These titles are referred to as **characteristics**.	The eight participants then gave a title to each group that best represented what the **descriptions** in that group meant. These titles are referred to as **characteristics**.
The variation across the levels of fluorosis in the number of participants who used each characteristic to describe the photographs was analysed using Chi-square.	The variation across the levels of fluorosis in the number of participants who endorsed each characteristic on the tick list when describing the photographs was analysed using Chi-square.

2.4.3. The combination of spontaneously generated words into groups in preparation for analysis

In all, the 90 participants spontaneously generated 170 descriptions when describing the photographs in the first stage of the study. These were subsequently collapsed into 32 groups since many of the descriptions were very similar in meaning, but were used too infrequently to be analysed separately. Allocation of the descriptions into groups was undertaken by eight further volunteers, who were unassociated with the main study.

A number of methodology books were read in order to find specific guidelines on how to group the descriptions (Bouma & Atkinson, 1995; Burgess, 1991; Cahoon, 1987; De Vaus, 2002; Hakim, 1982; Hakim, 1987; Hall & Hall, 1996; Patton, 1990; Strauss & Corbin, 1990). However, none of the books gave guidelines that were directly applicable to this investigation; instead they tended to give general rules of thumb. Hall & Hall (1996), for example, stated that:

> Open-ended questions... (require) some form of content analysis, in which the various responses to the question are grouped into a... set of discrete categories... This process is quite subjective... Where the words are similar but different, you have to exercise your judgement (p.133-4).

Therefore statistical advice was sought and the following method agreed. The eight participants were shown the complete list of descriptions spontaneously generated by the participants during Stage 1 of the study. They were asked to group these descriptions into categories of characteristics based on how similar they were in meaning.

After the eight participants had grouped the descriptions the level of agreement between them was investigated. The final grouping of the descriptions reflected the most common understanding among the participants. The criterion for a description being included in a certain group was that half of the participants placed it in that group. If the votes for any description were split evenly between two groups, or any description was not placed into the same characteristic group by at least half of the participants, then that description was classified as being miscellaneous.

The participants were also asked to choose one description that best represented the words in each group, e.g. one group may contain the words "clever", "wise", "intelligent", this group might then be given the title “intelligent”. The descriptions that were chosen to represent the groups will be referred to as **characteristics** from this point on. The participants were given the option of choosing their own characteristics to represent the descriptions or of grouping the descriptions under characteristics provided by the experimenter, or of using a combination of their own and the experimenter’s characteristics.

The characteristics and the descriptions that comprised them are shown in Table 2.1.

Table 2.1. Descriptions and their associated characteristics generated in Stage 1 of the study

Spontaneously generated descriptions	Characteristic
Charismatic; Attractive; Good looking; Actor/Model/Hollywood; Sexy	Attractive
Unattractive; Gawkish; Minging; Goofy; Teeth off-putting	Unattractive
Vain; Neat; Fashionable; Takes care of appearance; Presentable; Well kempt; Smart; Well groomed	Care Taken with Appearance
Untidy; Shabby; Does not take care of appearance; Unkempt; Scruffy clothing	Care Not Taken with Appearance
Sensible; Responsible; Methodical; Careful; Reserved; Serious; Fastidious; Thorough; Tidy; Organised	Careful
Careless; Neglectful; Irresponsible; Forgetful; Inattentive	Careless
Intelligent; Wise; Clever	Intelligent
Unintelligent; Dopey; Poor judgement; Silly; Short attention span; Stupid	Unintelligent
Hygienic; Clean; Particular with hygiene	Hygienic
Unhygienic; Bad smell; Dirty; Unclean	Unhygienic
Sporty; Healthy eater; Strong; Fit; Healthy; Active; Energetic	Healthy
Unhealthy; Ill; Smoker; Inertia	Unhealthy
Outgoing; Confident; Relaxed; Self-assured; Positive; In-control; Assertive; Not shy; Extrovert; Cocky	Confident
Insecure; Unconfident; Nervous; Self conscious; Shy	Insecure
Open; Frank; Honest; Trustworthy; Loyal	Honest
False; Unreliable	Dishonest
Diligent; Hard working; Purposeful; Ambitious; Successful; Reliable	Hard Working
Lazy	Not Hard Working
Kind; Generous; Understanding; Good listener; Warm; Caring; Kind hearted; Volunteer; Good; Helpful; Nice	Kind
Selfish; Cold hearted	Unkind
Happy; Optimistic; Cheerful; Joyful; Fun; Carefree	Happy
Depressed; Unhappy	Unhappy

Table 2.1 (continued). Descriptions and their associated characteristics generated in Stage 1 of the study

Spontaneously generated descriptions	Characteristic
Friendly; Bubbly; Fun loving; Easy to get along with; Sociable; Easy going; Likes to party; Polite; Enjoys himself/herself/life; Cheeky; Interested; Lively; Talkative; Smiley	Sociable
Bad mannered; Unapproachable; Aggressive; Loner; Trouble maker; Dull; Unsociable	Unsociable
Humorous; Funny; Good sense of humour	Humorous
Young; Childlike; Immature; Student	Young
Old	Old
Greedy; Wealthy; Higher social class; Glistening	Wealthy
Poor; Unfortunate; Lower social class	Poor
Promising; Down to earth; Enthusiastic; Pleasant	Positive miscellaneous words
Obsessive; Highly strung; Qualified womaniser; Boring	Negative miscellaneous words
Unintimidating; Average; Normal; Quiet; Decided; Ordinary; Loud; Blasé; Messy; Drinker; Approachable	Neutral or ambiguous miscellaneous words

When dividing the descriptions into groups there tended to be positive and negative groups that were polar-opposite in meaning, for example physically attractive and physically unattractive. The 90 participants who generated the descriptions had been asked to indicate whether they regarded each of the words as positive or negative in meaning (Section 2.3.4). There was very little variation among the participants' evaluations of the descriptions, and very little variation between the participants' and the experimenter's evaluations of the descriptions. A small number of the words were regarded as neutral or ambiguous in meaning by both the participants and the experimenter, these words were listed under a theme of ambiguity. However, dividing the descriptions into positive or negative characteristics was largely unproblematic and will not be referred to again.

2.4.4. The combination of the characteristics on the tick list into groups as preparation for analysis

In the second stage of the study participants were asked to look at the photographs again, but this time to endorse descriptions from the tick-list. In preparation for analysis these 50 descriptions were also collapsed into 22 characteristics by the same participants who reduced the spontaneously generated words into characteristics, using the same method (Table 2.2).

Table 2.2. Descriptions and their associated characteristics utilised in Stage 2 of the study

Descriptions on the tick list	Characteristic
Beautiful; Good looking	Attractive
Not good looking; Ugly	Unattractive
Clean	Clean
Dirty	Dirty
Systematic; Prompt; Efficient; Orderly	Careful
Careless; Unpunctual	Careless
Unhealthy	Unhealthy
Intelligent; Observant	Intelligent
Gullible; Unobservant	Unintelligent
Humorous; Relaxed	Happy
Unhappy	Unhappy
Sentimental; Understanding; Helpful; Kind; Thoughtful; Forgiving	Kind
Egotistical; Rude; Malicious; Phoney; Unkind; Cruel	Unkind
Co-operative; Exciting; Tolerant; Sociable; Gossipy; Witty	Sociable
Stubborn; Hostile	Unsociable
Frank	Honest
Cowardly	Dishonest
Independent	Independent
Reliable	Reliable
Excited	Positive miscellaneous words
Greedy	Negative miscellaneous words
Overcritical; Unemotional; Oversensitive; Unconventional	Neutral or ambiguous miscellaneous words

2.4.5. Stage 1 - analysis of spontaneously generated descriptions

Chi-square was used to analyse variation in the frequency with which each characteristic was used across levels of fluorosis.

For the participants who viewed the extra-oral photographs without being asked to pay particular attention to the mouth (Condition 1), four characteristics were used to described the images on at least 20 occasions and were analysed using the Chi-square test of association. These were: confident, happy, intelligent, and sociable (Table 2.3). However, none of these characteristics showed significant variation across the levels of fluorosis.

For the participants who viewed the extra-oral photographs and were asked to pay particular attention to the mouth (Condition 2), the characteristics reported on at least 20 occasions were: confident, happy, intelligent, kind, and sociable (Table 2.4). As in Condition 1 none of these characteristics showed significant variation across the levels of fluorosis.

When participants viewed the intra-oral photographs (Condition 3) they did not use any characteristic often enough to allow a Chi-square analysis.

Table 2.3. Characteristics spontaneously generated when describing extra-oral photographs of fluorosis of varying severity, without being asked to pay particular attention to the mouth

Characteristic	Total number of responses across levels of fluorosis	Chi-square	Significance level	Percentage of participants who endorsed theme for each level of fluorosis			
				No fluorosis	Mild fluorosis	Moderate fluorosis	Severe fluorosis
Confident	23	1.0	ns	20.0	23.3	20.0	13.3
Happy	49	2.6	ns	50.0	43.3	40.0	30.0
Intelligent	24	0.4	ns	23.3	20.0	16.7	20.0
Sociable	49	0.4	ns	43.3	40.0	43.3	36.7

Table 2.4. Characteristics spontaneously generated describing extra-oral photographs of fluorosis of varying severity, when asked to pay particular attention to the mouth

Characteristic	Total number of responses across levels of fluorosis	Chi-square	Significance level	Percentage of participants who endorsed theme for each level of fluorosis			
				No fluorosis	Mild fluorosis	Moderate fluorosis	Severe fluorosis
Confident	28	5.6	ns	33.3	20.0	30.0	10.0
Happy	43	1.8	ns	36.7	43.3	36.7	26.7
Intelligent	32	1.0	ns	33.3	26.7	23.3	23.3
Kind	24	0.4	ns	23.3	20.0	20.0	16.7
Sociable	48	2.8	ns	50.0	36.7	43.3	30.0

2.4.6. Stage 2 - analysis of descriptions endorsed on the tick list

The words on the tick list were collapsed into characteristics, and, as in Section 2.4.5, Chi-square was used to analyse the variation in the number of endorsements the characteristic received across the levels of fluorosis.

At least 20 endorsements were made for 10 characteristics in study Conditions 1 and 2 (extra-oral photographs, without and with cueing to teeth, respectively), and 13 characteristics were similarly endorsed for study Condition 3 (intra-oral photographs). These are shown in Tables 2.5-2.7. These tables show the total number of endorsements each characteristic received across the levels of fluorosis out of a possible 120 (30 participants per condition x four levels of fluorosis), the Chi-square value, the p-value, and the percentage of participants who endorsed each characteristic at each level of fluorosis.

Table 2.5. Characteristics endorsed from the tick-list when describing extra-oral photographs of fluorosis of varying severity, without being asked to pay particular attention to the mouth

Characteristic	Total number of responses across levels of fluorosis	Chi-square	Significance level	Percentage of participants who endorsed theme for each level of fluorosis			
				No fluorosis	Mild fluorosis	Moderate fluorosis	Severe fluorosis
Attractive	44	15.5	p<0.001	46.7	46.7	46.7	6.7
Careful	30	3.0	ns	30.0	30.0	26.7	13.3
Clean	58	21.5	p<0.001	70.0	56.7	53.3	13.3
Happy	65	0.6	ns	60.0	53.3	50.0	53.3
Honest	28	2.6	ns	23.3	30.0	26.7	13.3
Independent	22	2.0	ns	23.3	20.0	20.0	10.0
Intelligent	45	3.2	ns	46.7	33.0	43.3	26.7
Kind	82	5.1	ns	80.0	70.0	70.0	53.3
Reliable	41	1.3	ns	40.0	33.3	36.7	26.7
Sociable	87	5.5	ns	80.0	80.0	73.3	56.7

Table 2.6. Characteristics themes endorsed from the tick-list when describing extra-oral photographs of fluorosis of varying severity, when asked to pay particular attention to the mouth

Characteristic	Total number of responses across levels of fluorosis	Chi-square	Significance level	Percentage of participants who endorsed theme for each level of fluorosis			
				No fluorosis	Mild fluorosis	Moderate fluorosis	Severe fluorosis
Attractive	53	13.3	p<0.005	66.7	43.3	46.7	20.0
Careful	36	7.9	p<0.005	36.7	33.3	40.0	10.0
Clean	60	18.4	p<0.001	66.7	60.0	56.7	16.7
Happy	74	5.9	ns	76.7	63.3	70.0	36.7
Independent	28	4.5	ns	30.0	23.3	30.0	10.0
Intelligent	69	13.5	p<0.005	73.3	66.7	60.0	30.0
Kind	99	15.9	p<0.001	96.7	83.3	90.0	60.0
Reliable	62	20.4	p<0.001	66.7	56.7	66.7	16.7
Sociable	92	25.0	p<0.001	90.0	86.7	86.7	43.3
Unattractive	26	34.8	p<0.001	10.0	6.7	10.0	60.0

Table 2.7. Characteristics endorsed from the tick-list when describing intra-oral photographs of fluorosis of varying severity

Characteristic	Total number of responses across levels of fluorosis	Chi-square	Significance level	Percentage of participants who endorsed theme for each level of fluorosis			
				No fluorosis	Mild fluorosis	Moderate fluorosis	Severe fluorosis
Attractive	29	21.6	p<0.001	43.3	13.3	40.0	0.0
Careful	37	29.5	p<0.001	56.7	13.3	50.0	3.3
Careless	22	19.8	p<0.001	6.7	20.0	3.3	43.3
Clean	36	29.5	p<0.001	46.7	16.7	56.7	0.0
Dirty	21	43.6	p<0.001	3.3	10.0	0.0	56.7
Happy	30	13.0	p<0.005	33.3	16.7	43.3	6.7
Intelligent	26	10.0	p<0.005	33.3	13.3	33.3	6.7
Kind	43	8.2	p<0.005	50.0	33.3	43.3	16.7
Reliable	23	6.6	ns	26.7	10.0	30.0	10.0
Sociable	52	25.2	p<0.001	60.0	30.0	70.0	13.3
Unattractive	27	45.6	p<0.001	6.7	13.3	3.3	66.7
Unhealthy	27	69.3	p<0.001	0.0	13.3	0.0	76.7
Unkind	23	7.5	ns	3.3	23.3	20.0	30.0

As can be seen from Table 2.5, only two characteristics significantly varied across the levels of fluorosis when participants viewed extra-oral photographs without being cued to look at the mouth; these were attractive and clean. Endorsements for attractive remained constant at 46.7% across no, mild, and moderate fluorosis, before a drop of 40% to just 6.7% for severe fluorosis. Endorsements for clean were 70% for no fluorosis, dropping to 56.7% and 53.3% for mild and moderate fluorosis respectively, and then falling to 13.3% for severe fluorosis.

The number of characteristics that showed significant variation increased to eight when participants viewed extra-oral photographs and were asked to pay particular attention to the mouth (Table 2.6). As before, attractive and clean varied significantly, as did careful, intelligent, kind, reliable, sociable, and unattractive. The majority of the characteristics received fewer endorsements for severe fluorosis than for the other levels, but this trend was reversed for unattractive. Ten percent of participants rated the photographs as unattractive when they had normal enamel or moderate fluorosis, and just 6.7% rated them as unattractive when they had mild fluorosis, with endorsements for severe fluorosis then showing a steep climb to 60%. As might be expected, this pattern was reversed when the participants were endorsing the photographs as attractive. When rating photographs with normal enamel 66.7% of participants regarded them as attractive, 43.3% endorsed them as attractive with mild fluorosis, 46.7% with moderate fluorosis, and 20% with severe fluorosis.

Table 2.7 shows that 11 characteristics varied significantly in the endorsements participants gave them when looking at intra-oral photographs; these were attractive, careful, careless, clean, dirty, happy, intelligent, kind, sociable, unattractive, and unhealthy. As in the previous experimental Conditions there was a tendency for characteristics that might generally be considered to be positive to show fewer endorsements for severe fluorosis compared to the other levels, and for the converse to be true of negative characteristics. However, there was a trend when viewing intra-oral photographs that was not evident with extra-oral views: endorsements for mild fluorosis were closer to severe fluorosis than to moderate fluorosis and normal enamel for several of the characteristics. For example, endorsements for intelligent were 33.3% for no and moderate fluorosis and 13.3% and 6.7% for mild and severe fluorosis respectively.

When endorsing characteristics, participants made more attributions based on fluorosis in an extra-oral photograph when they were cued to look at the mouth than when they were not. Moreover, they made further attributions still when shown intra-oral photographs, which could be regarded as a further level of cueing to the mouth. In other words, there is a trend that indicates the greater the level of cueing to the mouth, the more attributions are made on the basis of dental fluorosis.

2.5. Discussion

The chief objective of this study was to identify descriptions and produce photographs of fluorosis for use in later studies. This was successfully achieved by the production of the photographs (Section 2.3.1) and the subsequent identification characteristics that varied, according to the level of fluorosis, in how frequently they were endorsed.

In addition to the photographs produced and characteristics identified there were three key findings of this study, these were: (i) no, mild, and moderate fluorosis were generally associated with positive characteristics, and severe fluorosis was generally associated negative characteristics, (ii) the greater the level of cueing to the teeth participants received, the more their endorsements of characteristics varied across the levels of fluorosis, and (iii) participants endorsements of characteristics varied more across the levels of fluorosis when asked to choose from a tick-list than did the characteristics they generated themselves.

It is understandable that participants' endorsements showed more difference between moderate and severe fluorosis than between normal enamel and mild or moderate fluorosis since, as fluorosis becomes severe it results in darker lesions on the teeth and pitting of the enamel. This deviation from normal enamel is far greater than the white striations and marks that typify mild and moderate fluorosis. As was noted in Section 1.3.4, the debate on the aesthetic impact of fluorosis has centred on whether or not mild fluorosis is a problem (Ellwood & O'Mullane, 1995; Hawley et al., 1996; Riordan, 1993a). There is little dispute in the literature that severe fluorosis is a cause for aesthetic concern, and the current results support previous findings (Alkhatib et al., 2004; Clark, 1995; Clark et al., 1993; Ellwood & O'Mullane, 1995; van Palenstein Helderman & Mkasabuni, 1993). However, the fact that mild and moderate fluorosis received a similar level of positive endorsements to normal enamel can be regarded as surprising in that, while previous studies have been inconclusive about the impact of mild fluorosis, they have typically found moderate fluorosis to be problematic (Ellwood & O'Mullane, 1995; Hawley et al., 1996; Riordan, 1993a).

The explanation for this discrepancy between the current findings and those of Ellwood & O'Mullane (1995) and Riordan (1993a) could be that those studies asked the public to rate the dental aesthetics of actual patients, in a face-to-face setting. There could have been variation between the patients' general appearance (physical attractiveness, height, weight, etc) and their dental appearance (size and shape of the teeth, gingival condition, etc) that influenced participants' judgements despite instructions to focus on the tooth colour. However, in the current study the only variation between the photographs participants viewed was the level of fluorosis; this finding can be taken as evidence that it was fluorosis, and not any other factor, that was the cause of the differing descriptions. Similarly, Hawley et al. (1996), asked participants to rate intra-oral images which were not standardised, allowing the possibility that variation between the images not related to fluorosis (tooth size, etc) may have impacted on participants' ratings.

Another possible reason the findings of this study differed from those of Hawley et al. (1996) and other studies that used intra-oral images (McKnight et al., 1998; McKnight et al., 1999) is that the intra-oral perspective may magnify the impact of fluorosis. McKnight et al. (1998) and McKnight et al. (1999) found that mild fluorosis was more of an aesthetic problem than other dental conditions using intra-oral images (these were standardised images in the case of McKnight et al., 1999). But in a typical, real life social interaction people do not see each others teeth in close-up with the lips retracted. It was for this reason that the present study used standardised full face images smiling in such a way that the teeth were exposed. However, it was recognised that there was an opportunity to investigate whether previous dental studies that used intra-oral images of fluorosis may have found exaggerated opinions of fluorosis, caused by using an intra-oral perspective. Therefore it was decided to use both life-sized facial images and intra-oral images. A further level of cueing was added (extra-oral image with cueing to the mouth) in order to look for a trend in the level of cueing.

It was not possible to investigate the effect of cueing on participants' spontaneous descriptions of the photographs in this study, since none varied according to level of fluorosis, regardless of the level of cueing participants received. When considering the effect that level of cueing had on participants' endorsements of descriptions of extra-oral photographs, participants who did not receive any cueing to the mouth (Condition 1), were of special interest as they simulated real life more closely than participants in the other conditions. In this circumstance the level of endorsement only varied significantly across levels of fluorosis for two characteristics (attractive and clean). In contrast, participants' endorsements varied across fluorosis for eight characteristics when viewing extra-oral photographs with an instruction to pay particular attention to the mouth, and for 11 characteristics when viewing intra-oral photographs (Tables 2.5-2.7).

These findings suggest that fluorosis is not necessarily a salient feature of a person's face unless specifically brought to the attention of others. The variation caused by the level of cueing could be of relevance when assessing past studies of lay people's perception of fluorosis that relied exclusively on intra-oral photographs, especially since the papers that have found mild fluorosis to be the greatest problem have relied on intra-oral images (McKnight et al., 1998; McKnight et al., 1999) whereas the papers that asked lay people to rate the dental appearance of patients found mild fluorosis to be more comparable to normal enamel (Ellwood & O'Mullane, 1995; Riordan, 1993a). However, the designs of these studies are not directly comparable since McKnight et al. (1998) and McKnight et al. (1999) asked participants to rate fluorosis against other dental conditions rather than simply rate each image separately. When Hawley et al. (1996) asked participants to rate intra-oral images separately, mild fluorosis was considered preferable to normal enamel, so the type of scoring used and the level of cueing may impact on how problematic the public consider fluorosis to be. A number of recent studies have began to look at the personal characteristics associated with other dental conditions (decay) by utilising standardised extra-oral images (Eli et al., 2001; Feng et al., 2001; Newton et al., 2003) which presumably suggests a growing awareness of the limitations of intra-oral images.

The tick list was included in the design of this study so that in the event that participants did not spontaneously generate sufficient descriptions for analysis there would still be enough data to identify characteristics relevant to fluorosis. Although participants did generate sufficient characteristics for analysis when they viewed extra-oral photographs, they showed far greater variation when endorsing characteristics rather than generating their own. This raises several possibilities. It may be that participants spontaneously generated thoughts and opinions that are most salient to them, and that they endorsed opinions that are less salient. In other words, the descriptions that the participants endorsed from the tick-list were weaker opinions than those they generated themselves. As none of their spontaneously generated descriptions varied significantly across the levels of fluorosis, this would suggest that participants in this study did not hold strong views about dental fluorosis. Another possibility is that participants' were reluctant to make judgements about others unless prompted, due to a fear of seeming impolite. It is also conceivable that participants found it difficult to articulate their own descriptions of the photographs; their opinions may have been stable across Stages 1 and 2, but were more clearly expressed during Stage 2 because they found endorsing characteristics easier than generating them. Future studies that attempt to identify the public's description of fluorosis (or any other conditions) may benefit from adding a closed question to their design, in the event that participants are unwilling or unable to express their views clearly.

The general trend throughout both the free response and prompted response (tick-list) parts of the study was for positive descriptions to be more strongly associated with no, mild, and moderate fluorosis than with severe fluorosis, and, in the prompted response stage of the study, negative descriptions were more commonly associated with severe fluorosis than other levels. However an interesting variation of this trend occurred when participants were endorsing words to describe intra-oral photographs of the teeth. Participants' endorsements when rating mild fluorosis were more similar to their ratings of severe fluorosis than moderate fluorosis or normal enamel for a number of characteristics. The likely explanation for this is that in the digitally manipulated photographs used in this study, mild fluorosis was represented by white striations along the teeth whereas moderate fluorosis was represented by a more uniform whitening of the tooth surface that was probably less conspicuous. The extent to which the public can discriminate between normal enamel and fluorosis in its milder forms has yet to be established, as is evident by the contradictory reports regarding perceptions of fluorosis (Hawley et al., 1996; Lalumandier & Rozier, 1998).

Several characteristics were not included in the Chi-square analysis because they were a few endorsements short of 20, the number required for a Chi-square, and these tended to be negative characteristics. However if a characteristic is endorsed once or twice when describing somebody with no, mild or moderate fluorosis, but used 16 or 17 times when describing the same person with severe fluorosis, this would suggest that the characteristic is relevant to fluorosis, even if the relevant data cannot be statistically analysed, Section 2.5.1 describes how this issue was resolved. There could be a number of reasons why these negative characteristics were not sufficiently endorsed to be included in the Chi-square analysis. For example, negative characteristics were largely associated with severe fluorosis and positive characteristics with no, mild, and moderate fluorosis. Given that there were three levels of fluorosis associated with positive descriptions, compared to one level associated with negative descriptions, it is understandable that there were more positive descriptions than negative descriptions that were suitable for Chi-square analysis. Another possible reason is that participants felt it would be impolite to endorse negative descriptions in the study. One final reason could be that since the characteristics relevant to fluorosis were endorsed from the tick list, a positive characteristic relevant to fluorosis may have been on the tick-list, but its negative alternative, although relevant, may have been absent (this could also be true of positive characteristics not on the list). The characteristics on the tick list covered a wide range of attributes and were systematically selected based on previous studies (Section 2.3.2), yet it may still have been possible that a small number of relevant characteristics were missing from the list.

2.5.1. Selection of characteristics for use in future studies

From the foregoing, it seems reasonable that the negative alternatives of positive characteristics identified as being relevant to fluorosis are included in future studies. The converse is also true, in that where negative characteristics have been identified as being relevant to fluorosis but their positive alternatives have not the positive characteristic should also be included. If the alternative characteristics are included with the 12 characteristics identified as being relevant to fluorosis, across all three experimental Conditions, in future studies then there are 18 characteristics identified as relevant (it is 18 rather than 24 because some of the opposites of the characteristics were already identified as relevant, e.g. attractive and unattractive).

For each characteristic used in the subsequent studies of this thesis there is a characteristic with the opposite meaning; i.e. there is a descriptive dimension that incorporates polar-opposite characteristics, such as attractive and unattractive. This dimension will be referred to as a **descriptive theme** or **theme**, from this point on, while the two poles of the themes will be referred to as **characteristics**. The 18 characteristics and the themes they relate to, that were identified as relevant to fluorosis and which are included in the studies described in Chapters 3 and 4 are shown in Table 2.8.

Table 2.8. The characteristics and descriptive themes identified as relevant to fluorosis regardless of level of cueing, and used in the studies described in Chapters 3 and 4

Positive characteristic	Negative characteristic	Descriptive theme
Attractive	Unattractive	Physical attractiveness
Careful	Careless	Carefulness
Clean	Dirty	Cleanliness
Happy	Unhappy	Happiness
Healthy	Unhealthy	Health
Intelligent	Unintelligent	Intelligence
Kind	Unkind	Kindness
Reliable	Unreliable	Reliability
Sociable	Unsociable	Sociability

2.6. Conclusion

The chief aim of this study was to produce images and characteristics to use in later studies, and this was successfully achieved. Additionally, the findings of this study suggest that neither mild nor moderate fluorosis are perceived less favourably than normal enamel when the perceiver is rating extra-oral photographs, although severe fluorosis does have a notable negative impact on judgements. Furthermore, the extent to which participants' endorsements of characteristics varied across levels of fluorosis when rating the images was impacted on by the extent to which they were cued to look at the mouth; with a greater focus leading to more variation attributable to fluorosis. This has implications for studies of the aesthetic importance of fluorosis which relied exclusively on intra-oral photographs. Subsequent experiments in this thesis were designed to allow further study of the importance of cueing while assessing characteristics attributed to individuals with fluorosis.

Having described the production of the materials necessary to measure attitudes to fluorosis in this chapter, the following chapters describe two experiments that investigated such attitudes.

Chapter 3. The Use of a Response Latency Measure to Assess Attitudes to Fluorosis

3.1. Introduction to the implicit measure of attitudes to fluorosis

This, the second study conducted in the course of this thesis, was designed to take an implicit measure of participants' attitudes to fluorosis, using a response latency technique. As described in Section 1.4, response latency measures have been used to prevent participants modifying their answers in order to conform with what they belief to be typical or desired social responses. Additionally, the use of a response latency technique may measure attitudes people are unaware they have. As noted previously (Section 1.4), attitudes can guide people's behaviour even when they are not actively engaged in deliberative thinking. Therefore, even if an individual is not consciously aware they have a given attitude it might influence their behaviour (Section 1.4). The aim of the study described in this chapter was therefore to investigate attitudes towards varying degrees of fluorosis (none, mild, moderate and severe), using a measure that utilised response latency.

The hypotheses of this study were as follows; firstly participants response directions will show more negative responses (i.e. "yes" responses to negative characteristics and "no" responses to positive characteristics) and fewer positive responses (i.e. "yes" responses to positive characteristics and "no" responses to negative characteristics) the greater the severity of the fluorosis. Secondly the more severe the level of fluorosis the participants viewed, the longer their response latency would be when they made positive responses, and the quicker their response latency would be when they made negative responses. These hypotheses are based on the rationale that normal enamel has been shown to be more positively regarded than mild fluorosis (in some studies), which in turn is regarded more positively than moderate fluorosis, and that all three are regarded more positively than severe fluorosis (Section 1.3.4). The predictions relating to response latency are based on the literature described in Section 1.4 where positive responses (such as the endorsement of a positive characteristic or rejection of a negative characteristic) are facilitated by positive stimuli and negative responses are facilitated by negative stimuli.

3.2. Method, materials and procedure

3.2.1. Standardised images used in this study

A total of 12 images were used in this study, they are described in Table 3.1. The images included the eight extra-oral photographs generated as described in Section 2.3.1 and shown in Appendix B. These images showed a standard male and female face with normal enamel, and the same faces digitally manipulated to show mild, moderate, and severe fluorosis; these will be referred to as the target images. In addition, four filler images were used; these were unaltered images of different subjects, photographed under identical conditions to the extra-oral images, as described in Section 2.3.1. The purpose of the filler images was to discourage participants from explicitly searching for variations in the target faces. No intra-oral images were used in this study.

Table 3.1. Images used to assess attitudes to fluorosis

Image	Type of image	Gender of photographed person	Level of fluorosis digitally added
1	Target	Male – same face in these images	None
2	Target		Mild
3	Target		Moderate
4	Target		Severe
5	Target	Female – same face in these images	None
6	Target		Mild
7	Target		Moderate
8	Target		Severe
9	Filler	Male	None
10	Filler	Male	None
11	Filler	Female	None
12	Filler	Female	None

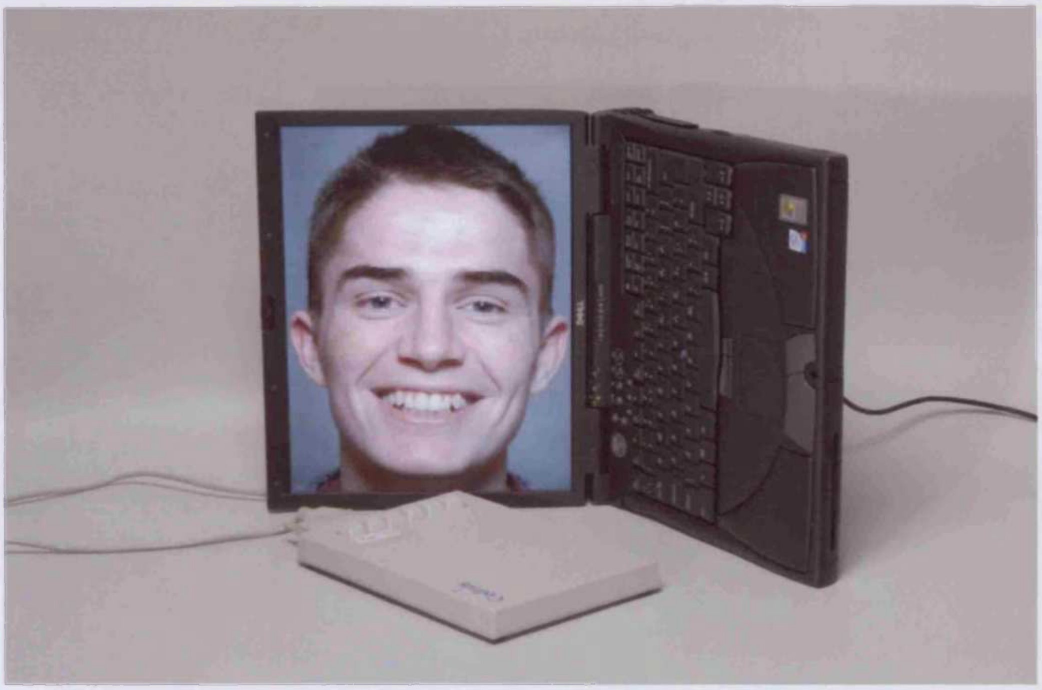
3.2.2. Characteristics

Eighteen characteristics were identified as being relevant to fluorosis in the previous study (Section 2.5.1) were selected for use in this study and are shown in Table 2.8, Section 2.5.1.

3.2.3. Computer hardware

The colour images and characteristics were displayed on a laptop computer, (Dell Notebook Inspiron 8200, Dell; Round Rock, Texas) the Liquid Crystal Display screen had a diagonal size of 382mm and a screen resolution of 1600 x 1200 pixels. A Cedrus 6-key serial response box RB-610 (Cedrus; San Pedro, CA) was used in conjunction with the computer. A photograph of the laptop and response box is shown in Figure 3.1.

Figure 3.1. Photograph of laptop and response box



3.2.4. Computer software

Cedrus SuperLab Pro 2.01 (Cedrus; San Pedro, CA), running on Microsoft Windows XP Pro operating system, was used to determine response latency. This priming software displays stimuli at programmed intervals and for programmed durations and also records responses made by the user, in terms of the length of time taken by respondents to depress the key following display of the stimulus, and which key on the response box is selected. Cedrus SuperLab software and equipment has been successfully used in previous response latency studies (Garcia & Bargh, 2003; Giner-Sorolla et al., 1999, Section 1.4.2.1) and is well established as a manufacturer of response latency equipment for use in psychological testing (E.T.S.L., 2001).

3.2.5. Development and programming of the computer based assessment

In order to measure response latency it was necessary firstly to programme the software to display the images and characteristics in the desired order. The programme was designed by the author and developed in collaboration with Mr Phil Heard, research technician in the Psychology Department of the University of Wales, Swansea. The laptop and software used to run the programme were also provided by this department, as was technical assistance.

3.2.6. Randomisation of images and characteristics for display

Participants' attitudes to fluorosis were gauged by viewing eight images of fluorosis, four filler images, and 18 characteristics. This resulted in 432 experimental trials per participant. Each trial involved the presentation of one image and one characteristic and is described in further detail in Section 3.2.9. The filler images were shown four times per characteristic so that the filler images and target images appeared with equal frequency (although level of fluorosis varied across the target images). This was done to avoid making the target faces seem of special relevance to the participants, and thus avoid motivating the participants to deliberately look for variations in the target images.

The 432 trials were randomised and divided into four equal blocks of 108 trials each, with one proviso; the target trials involving severe fluorosis were all placed into the final block and the other target trials were randomly dispersed throughout the first three blocks, the rationale for this will be discussed in Section 3.4. The randomisation was conducted using the randomisation function on Excel (Microsoft; Redmond, WA) and SPSS (SPSS Inc; Chicago, Illinois).

In addition to the target images showing no, mild, and moderate fluorosis, blocks 1-3 also contained two randomly selected filler faces per target face. The final block consisted of the target images showing severe fluorosis, plus one randomly selected filler image per target trial and one systematically selected target face showing no, mild or moderate fluorosis per severe fluorosis trial. The target images in the final block with no, mild, or moderate fluorosis were selected to give an even balance between gender and level of fluorosis; these images were regarded as filler images and responses to them were not analysed. The rationale for this will be discussed in Section 3.4. The type and number of images displayed in each block is shown in Table 3.2.

Table 3.2. Number of each type of image in each block

Block	No fluorosis male	No fluorosis female	Mild fluorosis male	Mild fluorosis female	Moderate fluorosis male	Moderate fluorosis female	Severe fluorosis male	Severe fluorosis female	Filler faces
1	7	7	8	7	7	9	0	0	63
2	5	7	5	4	4	4	0	0	79
3	6	4	5	7	7	5	0	0	74
4	6*	6*	6*	6*	6*	6*	18	18	36

*These target images were regarded as filler images and were not analysed.

Although the block in which each trial appeared was predetermined, the order in which trials appeared within that block was randomised by the Cedrus SuperLab Pro software; this was done to avoid order effects.

Altogether the filler trials (including the trials where target faces served as filler faces in block four) accounted for two thirds of the 432 trials, these were discarded when the data were analysed.

3.2.7. The question sheet

Participants were asked to complete a simple question sheet that was designed to ascertain whether or not they had noticed the differing dental appearance of some of the images displayed in the computer based assessment (Appendix D 8). This question sheet asked the participants if they had noticed anything unusual about any of the faces in the study but made no mention of dental appearance. It also asked them to give a rating of how confident they were that they had noticed something unusual on a seven-point scale, where 1 equalled very confident and 7 not at all confident.

The question sheet was administered twice, once during the three minute rest period between blocks three and four, and a second time after participants had completed the fourth and final block (Figure 3.2). This was done so that it would be possible to determine if mild and moderate fluorosis were explicitly noticed to the same extent as severe fluorosis.

3.2.8. Recruitment of participants

Contact was made with the administrative office of the Department of Biomedical Sciences. They e-mailed students on behalf of the experimenters requesting volunteers for the study (Appendix D 1). First year students on several courses (medicine, biology, zoology, biochemistry) and all second year bioscience students were e-mailed (in total 597 e-mails were sent). These e-mails were dispatched in stages in order to maintain a steady supply of volunteers. When the volunteers from one course were exhausted, the next course was e-mailed. First the medical students were e-mailed, then the biology, zoology, and biochemistry students, and finally the second year biosciences students. Students were e-mailed only once, if they did not respond to the first e-mail no reminders were sent.

Additionally posters were displayed in the Department of Biomedical Sciences and the Student Union Building requesting volunteers for the study and leaflets were handed out at the Arts and Social Sciences library in Cardiff. Arrangements were made, with students who wished to participate, to meet at mutually convenient time in one of the tutorial rooms in the Department of Biomedical Sciences building or in the Student Union Building, Park Place, Cardiff. All students were kept blind to the fact that the study was connected with dental appearance during the recruitment stage. It was for this reason a venue outside the dental school was selected.

3.2.9. Procedure

The programme took approximately one hour to run per participant, including the time taken for the instructions and a verbal debrief to be given, the full procedure is shown in Figure 3.2.

Figure 3.2. Procedure used to assess attitudes to fluorosis with response latency

Participants are assigned to either the cued or non-cued condition. They are given verbal and written instructions and give their demographic data and written consent.

↓

Participants begin computer administered assessment. In the first section they rate the valence of the characteristics.

↓

The practice block; participants are familiarised with the procedure using images and characteristics different from those used in the experimental trials. One minute break on completion of practice block.

↓

Experimental block 1, followed by a three minute rest period.

↓

Experimental block 2, followed by a three minute rest period.

↓

Experimental block 3, followed by a three minute rest period during which participants complete a question sheet to assess if they noticed anything unusual about the images.

↓

Experimental block 4.

↓

Participants complete the question sheet assessing if they noticed anything unusual about the images for a second time. Participants are then debriefed, reimbursed for travel expenses, and thanked. The procedure is complete.

Participants completed the study under one of two experimental conditions. They were, or were not, cued to look at the mouth of the images displayed. The instruction to pay close attention to the mouth was given verbally by the experimenter before participants began the experiment. It was repeated in an information sheet the participants were given (Appendices D 2 and D 3) and again in the on-screen instructions (Appendices D 6 and D 7). The two experimental conditions were identical apart from the presence or absence of the instruction to pay close attention to the mouth.

Before beginning the experiment participants were given the following verbal instructions by the PhD candidate, who remained present throughout the experiment:

> This is a study of first impressions based on facial appearance. We will show you a series of faces on the computer monitor, each face will stay on the screen for about two seconds, and then it will be replaced with a word. When you look at the faces will you please pay particular attention to the mouth. (This sentence was only included in the cued condition). We would like you to press the "yes" button if you think the word describes the face, and the "no" button if you think it doesn't. This is the main part of the study. However, before you begin this there is a very brief section where you will be shown words on the screen, one at a time. In this section we would like you to press the "yes" key if you think the word is positive in meaning, and the "no" key if you think it is negative in meaning. Please try to respond as quickly and accurately as you can in each trial. Do you have any questions?

After the verbal instructions and information sheet had been administered participants gave informed consent (Appendix D 4) and completed a demographic information sheet (Appendix D 5).

Participants were asked if they were right handed or left handed, and the response box was placed next to their dominant hand. The laptop was placed on its side so that the monitor could display the faces in a portrait view and participants began the computer administered task and question sheet assessment.

The participants' first task was to rate the valence of the characteristics used the in the practice and experimental trials. This was done by presenting characteristics one at a time on the monitor and asking participants to press the "yes" key if they believed the characteristic to have a positive meaning, and the "no" key if they believed it to have a negative meaning (two of the keys on the response box were labelled "yes" and "no"). The purpose of this was to confirm that participants had the same understanding of the characteristics as the experimenters.

After this they completed 20 practice trials, which consisted of images and characteristics that were not used in the main study, before undertaking the experimental trials. On screen instructions were presented before each of these stages.

The trials that formed the practice and experimental blocks involved five basic steps which are summarised in Figure 3.3.

Figure 3.3. Procedure of experimental trials in assessment of attitudes to fluorosis using response latency

Life-size, full face images, showing either normal enamel, mild, moderate, or severe fluorosis were displayed on a computer monitor for 2000ms. The face was then removed and the screen remained blank for 500ms, except for a cross to orientate participants to the next step.

↓

One of 18 characteristics identified as relevant to fluorosis appeared on the screen. These were presented in black, lower case lettering on a white background in Ariel bold font at 50 point.

↓

Participants pressed either a "yes" or "no" key on the response box to indicate whether or not they thought the characteristic described the previous face.

↓

The time the participant took to respond (in milliseconds) was recorded by the computer, as was the key depressed.

There was a rest period of 1 minute between the completion of the practice trials and commencement of the experimental trials, and there were rest periods of three minutes between each experimental block.

On completion of the study, participants were debriefed and paid £10 as reimbursement for travel expenses.

3.2.10. Ethical approval

The study was approved by the Bro-Taf local research and ethics committee on the 16^{th} April 2003.

3.2.11. The pilot study

A pilot study was undertaken to validate the equipment used in this study. Upon completion of the study the first 19 participants were given a simple question sheet designed to assess how easy it was to use the equipment and complete the study (Appendix D 9). The questions asked participants to rate their satisfaction with the response box and other equipment, the instructions, the colours used, the clarity of the words and faces, the display time of the words and faces, and the length of the whole study, and it also gave the participants the option to voice any other comments. These participants generally reported a high level of satisfaction with the equipment and procedure, therefore no changes were made to the study and the 19 participants were included in the main analysis.

3.2.12. Analysis

The direction of the analysis was guided by the PhD candidate but the appropriate statistics were suggested and models fitted by a statistician. These were primarily conducted using the software package MLwiN (Centre for Multilevel Modelling; London, U.K.) although SPSS was also used.

The data produced in this study had two outcomes, response direction (i.e. whether participants responded "yes" or "no") which is binary data, and response latency (the time participants took to respond in milliseconds) which is continuous data. Further to this, there were two levels to both types of data; response differences between the different participants, and within participant variation in responses. Therefore, the main analysis of Study 2 involved a two-level logistic regression to analyse response direction, and a two-level linear regression to analyse response latency. The levels of both these models are shown in Figure 3.4.

Figure 3.4. Two-level regression models

Level 2 – Between participant factors cueing, participant gender, participant age.

↓

Level 1 - Within participant factors Valence of characteristic, level of fluorosis, descriptive theme, response latency (linear regression only), response direction, gender of image

3.2.12.1. The logistic regression

The binary data consisted of "yes" and "no" responses; the dependent variable in the logistic regression was the likelihood of a "yes" response. The basic model was a two level logistic regression; this allowed for a random constant effect (i.e. individuals having differing underlying probabilities of saying yes). All other factors were fitted initially as fixed effects, with an assessment made as to whether a random effect would improve the model.

The basic model included two main effects, the first was the valence of the characteristic, and the second was the level of fluorosis, and it also looked at the interaction between these variables. After the variation caused by valence of characteristic and level of fluorosis had been attributed, the model was re-run with other variables added to it. The other variables were added in the following order:

- Level of cueing.
- The interaction between level of fluorosis and cueing.
- The interaction between cueing and the valence of the characteristic.
- The three-way interaction between level of fluorosis, valence of characteristic and cueing.
- Gender of image.
- Gender of participant.
- Age of participant.

If any of the variables did not account for significant additional variance, at the 5% level, in the model, they were discarded before the next variable was added. The coding levels used of all variables are shown in Table 3.3.

Table 3.3. The coding levels of variables in the logistic regression

Variable	Coding levels
Level of fluorosis	There were 3 dummy variables fl1, fl2, and fl3, scored as follows: fl1 = 0, fl2 = 0, fl3 = 0 = no fluorosis fl1 = 1, fl2 = 0, fl3 = 0 = mild fluorosis fl1 = 0, fl2 = 1, fl3 = 0 = moderate fluorosis fl1 = 0, fl2 = 0, fl3 = 1 = severe fluorosis
Valence of characteristic	1 = positive characteristic, 0 = negative characteristic
Cueing	1 = not cued, 0 =cued
Participant gender	1 = male, 0 = female

When the most suitable model was discovered using the entire data set this process was applied to an individual regression for each of the descriptive themes (pairs of characteristics describing polar-opposite characteristics).

3.2.12.2. The linear regression

The data produced by response latency studies is often negatively skewed and is therefore usually transformed before analysis (Fazio, 1990b; Fazio, 1993b). Both logistic and reciprocal transformation have been used (Fazio, 1993b; Greenwald et al., 1998). Therefore data were transformed using both of these methods and tests of normality were conducted and compared with each other and the raw data. The more effective transformation was then used in the subsequent analysis.

The linear regression model was designed to compare the (transformed) response latency of responses to each level of fluorosis to the response latency of responses to no fluorosis. As in the logistic regression, the model fitted was a two-level regression which allowed for a random constant effect. That allowed for individuals to have different mean response latencies although it was assumed that the distribution of their response latencies (within person) was not significantly different across individuals. The variables entered into the linear regression were decided on a similar basis to those in the logistic regression (Section 3.2.12.1). However, there were more variables in the linear regression. Therefore the variables were placed into groups based on their theoretical interest and entered one group at a time. The coding of each variable is shown in Table 3.4.

Table 3.4. The coding levels of variables in the linear regression

Variable	Coding levels
Level of fluorosis	There were 3 dummy variables fl1, fl2, and fl3, scored as follows: fl1 = 0, fl2 = 0, fl3 = 0 = no fluorosis fl1 = 1, fl2 = 0, fl3 = 0 = mild fluorosis fl1 = 0, fl2 = 1, fl3 = 0 = moderate fluorosis fl1 = 0, fl2 = 0, fl3 = 1 = severe fluorosis
Valence of characteristic	1 = positive characteristic, 0 = negative characteristic
Response direction	1 = yes, 0 = no
Cueing	1 = not cued, 0 =cued
Participant gender	1 = male, 0 = female

There were three groups of variables; the first group formed the basic model, and contained the variables for the main effects of interest plus one interaction, namely level of fluorosis, response direction, valence of characteristic, the interaction between response direction and valence of characteristic, and participant gender. The second group contained the interactions between level of fluorosis and the variables of interest, i.e. level of fluorosis with response direction, level of fluorosis with valence of characteristic, and level of fluorosis with response direction and valence of characteristic. The third and final group contained the variables related to level of cueing; these were the level of cueing variable, and the interaction between the level of fluorosis and level of cueing. The variables contained in each group are shown in Table 3.5. If the second and third groups did not explain any additional variance, at the 5% level, they were discarded.

Table 3.5. The variable groups added in the linear regression

Group	Variables
1	• Level of fluorosis • Response direction • Valence of characteristic • The interaction between response direction and valence of characteristic • Participant gender
2	• The interaction level of fluorosis with response direction • The interaction level of fluorosis with valence of characteristic • The interaction level of fluorosis with response direction and valence of characteristic
3	• Level of cueing • The interaction between level of fluorosis and level of cueing

This process was repeated for nine further linear regressions that isolated responses to descriptive themes. Participant gender was not included in the individual regressions because the imbalance in males and females meant that there would not be sufficient numbers in some cells to make the analysis meaningful.

3.2.12.3. Other analysis

In addition to the main analysis, participants' valence ratings (positive versus negative) of the characteristics were compared to the overall ratings generated by the first study (Table 2.8) using a Kappa test.

The descriptive statistics generated by participants' responses to the question sheets administered after blocks three and four were reported.

3.3. Results

3.3.1. Participants

In total 80 students enrolled at Cardiff University participated in this study. One of the participants was omitted from the analysis due to a high rate of responses (over 10%) that were too fast to be meaningful (under 300ms, see Section 3.3.2.1). This left 23 males and 56 females with a mean age of 20.48 (± 2.43) years-old and an age range of 18-32 years-old.

3.3.2. Preparation of data

3.3.2.1. The discarding of inappropriate responses

When dealing with response latency data it is common to discard or recode outliers, although exactly when and how this should be done is a "judgement call" (Fazio, 1990b, p.84). There are numerous examples of this practice, and many variations on exactly where cut off points should be; for example Arndt et al. (2002) used responses that were within a range of 200-2000ms, Hermans, Crombez, & Eelen (2000) used 200-1500ms, Hermans et al. (1998) and Hermans et al. (1994) used 250-1500ms, Klauer et al. (1997) used 2 standard deviations, Pratto & John (1991) used 300-1500ms, and Greenwald et al. (1998) used 300-3000ms in the IAT, which has since become the most widely used response latency measure.

The current study differs from previous studies that have utilised response latency in that participants' responses are not truly automatic; that is there is no definitive right or wrong answer. This allowed an element of deliberation in responses which predictably led to longer response latencies than in previous work. Therefore, while the lower limit could be set at a similar cut off point to previous work since these limits are based on when participants become aware of the stimulus, it was not appropriate to use the same upper limit because this would exclude too much of the data. Initially it was thought that an upper limit would prove unnecessary, however the data clearly showed some extreme outliers, in one instance a participant took over three minutes to respond, which would suggest that they were no longer focused on the study at that time and therefore make that particular response meaningless. Therefore an upper limit of 10,000ms was introduced, 99.8% of responses were within this limit. The lower limit was 300ms, based on previous work, and 0.7% of responses were below this limit, therefore 0.9% of all responses were discarded for being outside the range of 300-10,000ms.

Further to this some participants accidentally pressed one of the keys on the response box that was not the “yes” or "no” key and this was recorded as a response by the computer. These responses accounted for less than 0.1% of the total data set, just 14 out of 36,160 responses involved an invalid key response, and these were also discarded from the analysis.

Finally, based on Greenwald et al.'s (2003) recommendation that participants showing a high rate of erroneous responses be discarded from the data set, it was decided that any participant who responded outside the range of 300-10,000ms on 10% of his or her trials, or who pressed one of the unlabelled keys on 10% or more of his or trials would be excluded from the analysis. As described in Section 3.3.1 one participant’s data was discarded from the analysis because he responded faster than 300ms in over 10% of the trials which suggested random responding since so many responses were too quick to be meaningful.

Outliers were discarded, rather than recoded to the upper or lower limit because there were two response outcomes in the present study, response direction and response latency. Therefore if a response was regarded as being too fast or slow to be meaningful then the response direction (whether the participant said “yes” or “no”) also needed to be discarded.

After the outliers had been discarded the practice trials and filler trials were filtered from the data set before it was checked for normality.

3.3.2.2. The distribution of the response latency data

The data from the second study was negatively skewed, as is often the case with response latency data. Therefore it is typical to transform data before analysis (Fazio, 1990b; Fazio, 1993b). The data in this study were transformed using both logistic (log_{10}) and reciprocal transformation and then the transformations were compared for normality of distribution. The tests of normality conducted were skewness and kurtosis, and histograms were plotted. However, tests of skewness and kurtosis may be too sensitive when dealing with large samples and therefore Tabachnick & Fidell (2001, p.74-5) recommend the shape of the distribution be used for judging normality of distribution rather than statistics when dealing with such a sample. The current data set can be considered to constitute a large sample since it had 11340 cases provided by 79 participants; according to Tabachnick & Fidell (2001) 100 cases can be considered large in terms of positive kurtosis, and 200 cases is large in terms of negative kurtosis. The histograms and tests of normality were run twice, once on the entire data set, and once after the data set had been split by characteristic (even after the file was split by characteristic there were over 600 cases per test). The results of these tests are shown in Table 3.6 and in Figure 3.5.

The Kolmogorov-Smirnov test and the Shapiro-Wilk test were also performed. However, these are not reported since both tests showed non-significance for both types of transformation in almost every test (except for the Kolmogorov-Smirnov test for the characteristic “dirty” in the reciprocal transformation). Moreover, as was noted by Pallant (2001, p.58) when discussing the Kolmogorov-Smirnov test, these tests are also dependent on sample size, and in a large sample such as the one in this study they tend to be too sensitive to be meaningful.

As can be seen from Table 3.6 and the graphs in Figure 3.5, the majority of the characteristics are closer to normal distribution with logistic transformation, therefore the response latency data were logistically transformed before analysis.

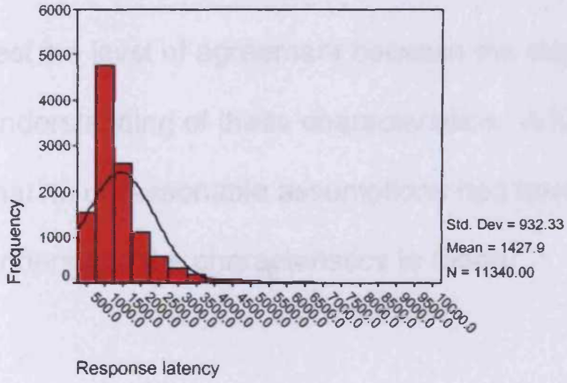
Table 3.6. Tests of normality of distribution after logistic and reciprocal transformation

Characteristic	Skewness - Transformation method closest to 0 (i.e. closest to normal distribution) shown in **bold**		Kurtosis - Transformation method closest to 0 (i.e. closest to normal distribution) shown in **bold**		Which Graph looked more normally distributed
	Logistic	Reciprocal	Logistic	Reciprocal	
All characteristics	**0.671**	0.951	**0.766**	1.724	Logistic
Attractive	**0.687**	0.857	**0.809**	1.972	Logistic
Careful	**0.386**	1.461	**0.412**	4.292	Logistic
Careless	**0.693**	0.819	**0.834**	1.128	No noticeable difference
Clean	**0.536**	1.077	**0.777**	2.581	Logistic
Dirty	1.165	**0.481**	2.119	**1.059**	Logistic
Happy	**0.686**	0.742	0.974	**0.652**	No noticeable difference
Healthy	0.792	**0.789**	**1.076**	1.465	Logistic
Intelligent	**0.556**	1.034	**0.251**	2.643	Logistic
Kind	0.805	**0.736**	1.274	**0.826**	No noticeable difference
Reliable	**0.660**	0.821	**0.372**	1.513	Logistic
Social	**0.763**	0.801	**1.005**	1.294	No noticeable difference
Unattractive	**0.463**	1.232	**0.584**	2.999	Logistic
Unhappy	**0.679**	0.912	**0.747**	1.661	Logistic
Unhealthy	0.689	**0.626**	0.648	**0.350**	No noticeable difference
Unintelligent	**0.591**	0.908	**0.635**	1.525	No noticeable difference
Unkind	0.784	**0.761**	**1.223**	1.242	No noticeable difference
Unreliable	**0.683**	1.168	**0.960**	2.319	Logistic
Unsociable	**0.565**	1.236	**0.531**	3.119	Logistic

Figure 3.5. Histograms of the data distribution for the entire data set before transformation and after logistic and reciprocal transformation.

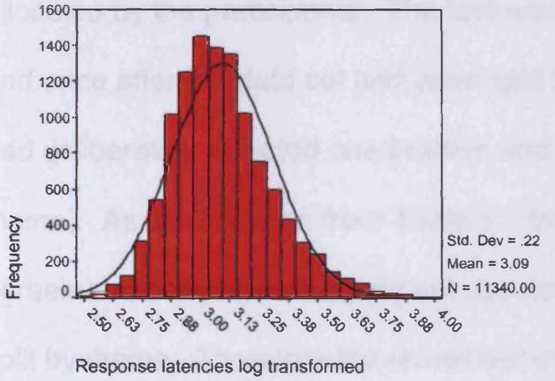
Complete data set (all characteristics)

Untransformed data



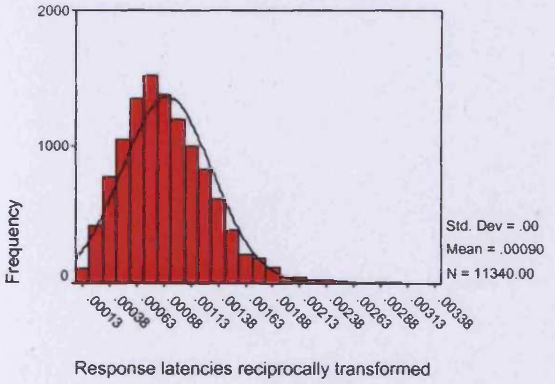
Complete data set (all characteristics)

Logistically transformed data



Complete data set (all characteristics)

Reciprocally transformed data



3.3.3. Participants valence ratings of the characteristics

The participants' first task in the experimental procedure was to rate each of the characteristics as being either positive or negative. The purpose of this was to test the level of agreement between the experimenters and participants' understanding of these characteristics. A high level of agreement would suggest that no unreasonable assumptions had been made about what participants understood the characteristics to mean.

The Kappa test was used to determine the level of agreement between the valence allocated to each characteristic by the experimenter and the valence allocated by the participants. The test was run twice, once on the entire data set, and once after the data set had been split by descriptive theme (the experimenter had deliberately selected one positive and one negative characteristic for each theme). As can be seen from Table 3.7 there was a very high level of agreement, both when the data set was considered in its entirety and when it was split by theme. Therefore the remainder of the analysis could be conducted in the knowledge that participants shared the valence ratings allocated to each characteristic by the experimenter.

Table 3.7. Agreement levels between experimenter and participant ratings of the valence of the characteristics

Theme	Kappa statistic	Significance
All themes combined	0.957	p<0.001
Attractiveness	0.949	p<0.001
Carefulness	0.936	p<0.001
Cleanliness	0.962	p<0.001
Happiness	0.975	p<0.001
Health	0.975	p<0.001
Intelligence	0.949	p<0.001
Kindness	0.987	p<0.001
Reliability	0.948	p<0.001
Sociability	0.936	p<0.001

3.3.4. Analysis of participants' response direction

In the main experimental procedure participants pressed a "yes" or "no" key to show whether they thought a characteristic described a face. The percentage of "yes" responses to each characteristic, for each level of fluorosis are shown in Table 3.8.

Table 3.8. Percentage of "yes" responses [a] to each characteristic by level of fluorosis

Characteristic	N [b]	Level of fluorosis			
		No	Mild	Moderate	Severe
Attractive	629	41.4	31.6	38.9	8.9
Unattractive	631	38.6	36.9	36.7	78.5
Careful	630	56.3	62.0	61.8	31.2
Careless	628	24.2	28.2	31.0	64.3
Clean	628	78.3	78.3	74.5	16.6
Dirty	630	16.5	9.5	15.3	81.5
Happy	630	85.4	89.7	88.0	74.7
Unhappy	631	11.4	10.8	15.8	15.9
Healthy	632	74.7	74.1	72.8	26.6
Unhealthy	630	20.3	20.9	25.9	73.1
Intelligent	629	66.9	66.9	71.3	46.8
Unintelligent	628	21.7	18.6	20.3	34.4
Kind	632	86.7	85.4	84.2	75.3
Unkind	630	7.6	7.0	8.3	15.8
Reliable	629	72.2	71.3	71.3	47.1
Unreliable	631	20.3	20.3	17.8	43.0
Social	630	86.0	85.4	83.4	58.9
Unsocial	632	9.5	13.3	11.4	33.5

[a] The remaining percentage are "no" responses.
[b] N varies slightly because responses considered erroneous were omitted.

Participants' response direction was analysed using logistic regression where the dependent variable was the likelihood of a "yes" response (the logistic regression equation is shown in Appendix E 1). Table 3.9 summarises the logistic regression conducted on the entire data set (i.e. the data was not split by descriptive theme).

Table 3.9. Two-level logistic regression model of response direction for all characteristics combined

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	20.99	(10.01, 44.03)	p<0.001
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.99	(0.74, 1.32)	
Moderate	1.14	(0.86, 1.52)	
Severe	6.17	(4.74, 8.04)	
Valence of characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.02	(0.70, 1.50)	
Moderate	0.81	(0.56, 1.18)	
Severe	0.02	(0.02, 0.04)	
Cueing	1.07	(0.59, 1.94)	ns
Valence of characteristic* Cueing	0.85	(0.30, 2.40)	ns
Cueing*Level of fluorosis			ns
No	1 [a]	R	
Mild	0.93	(0.62, 1.40)	
Moderate	0.95	(0.64, 1.43)	
Severe	0.89	(0.61, 1.30)	
Valence of characteristic* Cueing* Level of fluorosis			p<0.01
No	1 [a]	R	
Mild	1.00	(0.59, 1.72)	
Moderate	1.19	(0.70, 2.03)	
Severe	2.24	(1.34, 3.73)	

[a] No fluorosis is the reference category

Table 3.9 shows that the main effects of valence of characteristic and level of fluorosis were both significant, as was the interaction between them. The three-way interaction between valence of characteristic, level of fluorosis, and level of cueing (i.e. whether or nor participants were asked to pay particular attention to the mouth) was also significant, but the main effect of cueing and the two-way interactions between cueing and valence of characteristic, and cueing and level of fluorosis were not significant. Participants were more likely to give a "yes" response to a positive characteristic than to a negative characteristic, and were more likely to give a "yes" response to severe fluorosis than to no, mild, or moderate fluorosis. The interaction between level of fluorosis and valence of characteristic shows that participants were significantly less likely to give a "yes" response to a positive characteristic for severe fluorosis. This suggests that although "yes" responses were more common for severe fluorosis than for no, mild, or moderate fluorosis this can be largely attributed to the fact there were far more "yes" responses to trials involving negative characteristics when judging severe fluorosis than when judging the other levels of fluorosis. The three-way interaction shows that participants who were not cued to look at the mouth were significantly more likely to give a "yes" response to a positive characteristic when viewing severe fluorosis than cued participants, although this was not the case when viewing mild or moderate fluorosis.

The main effects of gender of image, participant gender, and participant age were also calculated but did not have a significant effect. Therefore, these variables were not included in the final model.

Table 3.10 is a prediction table created using the logistic regression conducted on the entire data set. This table clearly illustrates that, in comparison with normal enamel, or mild or moderate fluorosis, “yes” responses for severe fluorosis were far less likely for positive characteristics, but far more likely for negative characteristics, and that there was very little difference between the responses to no, mild and moderate fluorosis. This suggests that mild and moderate fluorosis either were not noticed by the participants, or were not considered to be a problem. However, when viewing severe fluorosis, participants clearly felt that the dental appearance differed from the norm so extremely that it was justifiable to make negative social judgments and to withhold positive social judgements. Table 3.10 also shows that uncued participants were more likely to give a “yes” response to a positive characteristic when viewing severe fluorosis than uncued participants, but other than this cueing had very little impact on participants’ direction of response.

Table 3.10. Prediction table for likelihood of a "yes" response to positive and negative characteristics by level of fluorosis and level of cueing

	No fluorosis		Mild fluorosis		Moderate fluorosis		Severe fluorosis	
	Negative characteristics	Positive characteristics	Negative characteristics	Positive characteristics	Negative characteristics	Positive characteristics	Negative characteristics	Positive characteristics
Uncued participants	15	75	14	74	16	76	48	48
Cued participants	14	77	14	77	15	76	49	34

Logistic regressions were conducted on the data after it was split into individual descriptive themes. The p-values of the main effects and interactions in these regressions are summarised in Table 3.11 (the logistic regression tables for the individual descriptive themes are reported in Appendix E 5). Where there were significant findings in these regressions they were in the same direction as those described under the analysis of the entire data set. The main effect of cueing and the interactions involving cueing were not significant in the majority of the regressions, and therefore were not fitted as part of the final model in these instances; these are indicated by empty cells.

Participants' social judgements on the themes of physical attractiveness, carefulness, cleanliness, health, intelligence, kindness, reliability, and sociability all showed the trend for severe fluorosis to be judged less favourably than normal enamel, such that, for example, participants would be less likely to say that a person was attractive if they had severe fluorosis (compared to normal enamel) and more likely to say that they were unattractive. Mild and moderate fluorosis were not shown to be different to normal enamel in any of these judgements. The only judgements that did not reveal a significant main effect of level of fluorosis were those judgements related to how happy a person appeared to be. However, the significant interaction between valence of characteristic and level of fluorosis indicates that the images with severe fluorosis were judged less favourably than images with normal enamel.

These findings suggest that the social judgements made about an individual are not altered if that individual has mild or moderate fluorosis, but do change unfavourably if they have severe fluorosis; and that this is true of social judgements made about a range of characteristics. The interaction between level of fluorosis and cueing only reached significance for two themes, kindness and reliability, suggesting cueing was not important in the majority of participants' judgements.

Table 3.11. Summary of p-values for each variable in the individual logistic regressions [a]

Regression analysis	Valence of characteristic	Level of fluorosis	Valence of characteristic* Level of fluorosis	Cueing	Cueing* Level of fluorosis	Valence of characteristic* Cueing	Valence of characteristic *Cueing* Level of fluorosis
All themes combined	$p<0.001$	$p<0.001$	$p<0.001$	ns	ns	ns	$p<0.01$
Attractiveness	ns	$p<0.001$	$p<0.001$	-	-	-	-
Carefulness	$p<0.001$	$p<0.001$	$p<0.001$	-	-	-	-
Cleanliness	$p<0.001$	$p<0.001$	$p<0.001$	-	-	-	-
Happiness	$p<0.001$	ns	$p<0.001$	-	-	-	-
Health	$p<0.001$	$p<0.001$	$p<0.001$	-	-	-	-
Intelligence	$p<0.001$	$p<0.001$	$p<0.001$	-	-	-	-
Kindness	$p<0.001$	$p<0.05$	$p<0.001$	ns	$p<0.05$	-	-
Reliability	$p<0.001$	$p<0.001$	$p<0.001$	ns	$p<0.01$	-	-
Sociability	$p<0.001$	$p<0.001$	$p<0.001$	-	-	-	-

[a] Some of the cells are empty because not all of the variables were fitted in every regression, see Section 3.2.12.1.

3.3.5. Analysis of participants' response latencies

In addition to participants response direction (whether they pressed the "yes" or "no" key) their response latency (the time between the presentation of the characteristic and the participant pressing a key) was recorded. This was subsequently analysed separately from response direction (although response direction was included as an explanatory variable in the analysis). This Section (3.3.5) describes the analysis of participants' response latency.

Descriptive statistics of the raw response latencies for the entire data set are shown in Table 3.12. The descriptive statistics include the median and the range because these are appropriate for skewed data. Furthermore, previous response latency studies have reported the mean and therefore the mean is included in Table 3.12 since it may be of interest for comparison to previous work. Tables showing the descriptive statistics of the raw response latencies for responses to each individual characteristic are shown in Appendices E 3 and E 4. Data for "yes" responses and "no" responses are shown separately because comparisons between the same response direction are more meaningful and informative since they show the time needed to make the same decision for different levels of fluorosis; a different response direction means a different decision, and possibly a different thought process.

Table 3.12. Response latencies (milliseconds), by level of fluorosis for all characteristics combined

Valence of characteristic	Response direction	No fluorosis			Mild fluorosis			Moderate fluorosis			Severe fluorosis		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Positive	Yes	1295	1091	370-9874	1299	1092	400-7040	1276	1051	331-9804	1614	1252	341-9474
	No	1606	1312	301-6710	1564	1297	341-7912	1535	1311	330-5107	1406	1162	341-7120
Negative	Yes	1767	1467	441-8262	1789	1467	451-8923	1923	1553	401-8913	1609	1332	501-8633
	No	1292	1041	320-9934	1288	1061	320-8713	1344	1092	341-9744	1592	1272	331-9584

The purpose of recording response latency was to measure the strength of participants' attitudes (with shorter response latencies indicating stronger opinions). This was particularly valuable when considering whether attitudes to mild and moderate fluorosis were different to attitudes expressed for normal enamel, since while there appeared to be no difference in the direction of the opinion (i.e. mild and moderate fluorosis were considered positive when normal enamel was considered positive, and considered negative when normal enamel was considered negative) there may have been a difference in the strength of the opinion (was normal enamel considered more positive or less negative than mild and moderate fluorosis?).

The response latency data was analysed using linear regression where the dependent variable was the transformed (to the $\log_{10}$) response latency. The final model selected for the linear regression is shown in Appendix E 2. The regression table for this equation is Table 3.13, and the impact of the main effects and interactions of interest are summarised in Tables 3.14 and 3.15.

Table 3.13. Two-level linear regression model showing the differences in response latency for all characteristics combined

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			p<0.001
No	0[a]	R	
Mild	0.001	(-0.016, 0.019)	
Moderate	0.011	(-0.007, 0.028)	
Severe	0.084	(0.065, 0.103)	
Response direction	0.095	(0.071, 0.119)	p<0.001
Valence of characteristic	0.063	(0.042, 0.083)	p<0.001
Valence of characteristic* Response direction	-0.153	(-0.186, -0.159)	p<0.001
Participant gender	0.104	(0.050, 0.159)	p<0.001
Level of fluorosis* Valence of characteristic			ns
No	0[a]	R	
Mild	0.003	(-0.015, 0.021)	
Moderate	0.010	(-0.009, 0.028)	
Severe	-0.001	(-0.020, 0.018)	
Level of fluorosis *Response direction			p<0.001
No	0[a]	R	
Mild	0.006	(-0.028, 0.040)	
Moderate	0.017	(-0.016, 0.050)	
Severe	-0.105	(-0.135, -0.074)	
Level of fluorosis* Valence of characteristic *Response direction			p<0.001
No	0[a]	R	
Mild	0.004	(-0.040, 0.049)	
Moderate	-0.017	(-0.061, 0.027)	
Severe	0.226	(0.185, 0.268)	

[a] No fluorosis is the reference category

Table 3.13 (continued). Two-level linear regression model showing the differences in response latency for all characteristics combined

	Coefficient	95% Confidence Interval	Significance
Cueing	0.027	(-0.021, 0.080)	ns
Cueing* Level of fluorosis			p<0.001
No	0 [a]	R	
Mild	-0.009	(-0.038, 0.019)	
Moderate	-0.023	(-0.051, 0.006)	
Severe	-0.124	(-0.151, -0.097)	

[a] No fluorosis is the reference category

As can be seen from Table 3.13 level of fluorosis has a significant effect on response latency, and, based on the coefficients and confidence intervals, it would appear that this difference is attributable to judgements involving severe fluorosis being slower than judgements involving the other levels of fluorosis. The precise nature of the difference is illuminated in Tables 3.14 and 3.15.

Table 3.14. Change in logged response latency (milliseconds) associated with level of fluorosis, valence of characteristic and response direction

Response direction	Level of fluorosis by valence of characteristic							
	No		Mild		Moderate		Severe	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
No	0	0.063	0.001	0.067	0.011	0.084	0.084	0.146
Yes	0.095	0.005	0.102	0.020	0.123	0.037	0.074	0.293

Table 3.14 shows the impact on logged response latency of response direction and valence of characteristic across levels of fluorosis. The basic model of the linear regression showed response latencies for "no" responses to negative characteristic, for normal enamel, therefore the value in this cell is 0. When judging faces with normal enamel or with mild or moderate fluorosis the fastest responses (i.e. the smallest numbers) were for "no" responses to negative characteristics. "Yes" responses to positive characteristic were only slightly slower, and then there was a larger increase in the time taken for "no" responses to positive characteristic, and a slower response still for "yes" responses to negative characteristic. In other words people were quickest to say something negative was not true about a person, and were almost as quick to say that something positive was true of a person. They took a longer time when saying that a person did not appear to possess a positive attribute, and their slowest responses were when endorsing something negative about a person. This pattern suggests that they held stronger opinions when making a complimentary response than when making a critical response. This may reflect social norms of politeness, in that participants were more hesitant to say something uncomplimentary about somebody than to say something positive about them.

While this trend was consistent across normal enamel, mild fluorosis and moderate fluorosis the actual response latencies themselves were slightly longer for mild fluorosis than for normal enamel, and slightly longer again for moderate fluorosis than for mild fluorosis. Although the confidence intervals in Table 3.13

show that these time differences were not statistically significant they may reflect a slightly longer period of evaluation caused by an increasingly unusual appearance of the teeth.

There was a different pattern in the response latencies to severe fluorosis. The fastest responses to severe fluorosis were for “yes” responses to negative characteristics, whereas this was the slowest type of response across the other three levels of fluorosis. “No” responses to negative characteristic were the next quickest type of response, followed by “no” responses to positive characteristic, and finally “yes” responses to positive characteristic, which were far slower than any other type of response across all four levels of fluorosis. This trend suggests that participants had strong attitudes when they were saying something negative about a person with severe fluorosis, but weak attitudes when saying something positive about them, which may indicate that social norms of politeness are suspended when somebody has a dental appearance that is obviously different to the norm.

As can be seen from Table 3.13 the main effect of cueing did not have a significant impact on logged response latency. However, the interaction between level of fluorosis and level of cueing did show a significant effect, the impact cueing had across levels of fluorosis is illustrated in Table 3.15.

Table 3.15. Change in logged response latency (milliseconds) associated with level of fluorosis and level of cueing

Level of cueing	Level of fluorosis			
	No	Mild	Moderate	Severe
Cued	0	0.001	0.011	0.084
Not cued	0.027	-0.008	-0.012	-0.04

Table 3.15 shows that the basic model was set up to show response latencies for cued participants judging photographs with normal enamel; therefore the value in this cell is 0. As can be seen from this table cued participants took a longer time than uncued participants to respond to severe fluorosis. Although there are differences between cued and uncued participants response latencies to the other levels of fluorosis the confidence intervals in Table 3.13 show these are not significant.

Further linear regressions were conducted on the data after it was split into individual descriptive themes. The p-values of the main effects and interactions of these regressions are summarised in Table 3.16 (the linear regression tables for the individual descriptive themes are reported in Appendix E 6). Where there were significant findings in these regressions they were in the same direction as those described under the analysis of the entire data set. As in the model fitted to the entire data set, variables were added in groups to see if they explained significantly more of the variance (see Section 3.2.12.2 and Table 3.5), for this reason not all of the descriptive themes have p-values for each variable.

Table 3.16. Summary of p-values for each variable in the individual linear regressions [a]

Regression analysis	Level of fluorosis	Response direction	Valence of characteristic	Valence of characteristic* Response direction	Subject gender	Level of fluorosis* Valence of characteristic	Level of fluorosis* Response direction	Level of fluorosis* Valence of characteristic *Response direction	Cueing	Cueing* Level of fluorosis
All themes combined	p<0.001	p<0.001	p<0.001	p<0.001	p<0.001	ns	p<0.001	p<0.001	ns	p<0.001
Attractiveness	p<0.01	p<0.001	ns	p<0.01	-	-	-	-	-	p<0.05
Carefulness	ns	ns	p<0.001	p<0.05	-	ns	p<0.01	p<0.001	ns	p<0.001
Cleanliness	ns	p<0.001	p<0.001	p<0.001	-	ns	ns	ns	ns	p<0.001
Happiness	p<0.001	ns	p<0.001	p<0.001	-	ns	p<0.001	p<0.001	p<0.05	p<0.001
Health	ns	p<0.01	ns	p<0.001	-	ns	p<0.001	p<0.001	ns	p<0.001
Intelligence	p<0.001	p<0.001	ns	p<0.001	-	-	-	-	-	-
Kindness	p<0.001	p<0.01	p<0.05	p<0.01	-	-	-	-	p<0.05	ns
Reliability	p<0.001	p<0.001	ns	p<0.01	-	ns	p<0.01	p<0.001	p<0.05	p<0.001
Sociability	p<0.001	p<0.001	ns	p<0.001	-	-	-	-	p<0.05	ns

[a] Some of the cells are empty because not all of the variables were fitted in every regression, see Section 3.2.12.2.

The three-way interaction between level of fluorosis, response direction, and valence of characteristic described above reached significance for the themes of carefulness, happiness, health, and reliability. Although this interaction was not statistically significant in the other themes it is likely it occurred to some degree since the when the entire data set was analysed it was highly significant ($p<0.001$), which would be unlikely if the trend only occurred in four of the nine themes.

The linear regressions conducted on the data after it was split into individual descriptive themes showed the main effect of cueing occasionally reached significance (for the themes of happiness, kindness, reliability, and sociability), and the interaction between cueing and level of fluorosis reached significance for the themes of physical attractiveness, carefulness, cleanliness, happiness, health, and reliability, which suggests that cueing participants had a particular impact on their response latencies to these themes.

In addition to the main effects and interactions involving level of fluorosis, response direction, valence of characteristic, and level of cueing, the linear regression also showed that participants' gender showed a significant effect such that males responded more slowly than females. However, the implications of this finding were tempered by the imbalance of males and females in the study. The main effects of gender of image and participant age were also calculated as part of the first group of variables but did not have a significant effect. Therefore, these variables were not included in the final model.

3.3.6. Assessment of participants' awareness of the varying dental conditions

The results of the question sheet that assessed how aware participants were of the varying levels of fluorosis showed that six out of 40 uncued participants noticed the teeth were changing before severe fluorosis was included, but all 40 noticed the teeth were changing after severe fluorosis was included. The results were very similar for cued participants, with seven out of 39 cued participants noticing the teeth were changing before severe fluorosis was included and all 39 noticing after severe fluorosis was included.

Participants gave a rating of how confident they were that they had noticed something unusual on a seven-point scale where 1 equalled very confident and 7 not at all confident. The mean confidence rating of the six uncued participants' who had noticed the altering dental appearance was 4.83, after viewing just no, mild, and moderate fluorosis, and 2.17 after seeing severe fluorosis. The mean confidence rating of the seven cued participants' who had noticed the altering dental appearance was 4.86, after viewing just no, mild, and moderate fluorosis, and 1.29 after seeing severe fluorosis. Therefore it appears that severe fluorosis allowed participants to feel more confident that the teeth they viewed were of abnormal appearance.

Severe fluorosis clearly had a very large impact on participants' explicit awareness of the changing dental appearance. However, cueing participants to the mouth appears to have had very little impact on how aware they were of the different levels of fluorosis.

3.4. Discussion

The main findings of this study are, firstly, that as in the previous study (Chapter 2), mild and moderate fluorosis do not appear to elicit different social judgements to normal enamel but severe fluorosis did have a negative impact on social judgements, when viewed in a standardised extra-oral photograph. This is true both in terms of the direction of the judgement and the strength of the attitude (as measured by response latency). This finding means that the experimental hypotheses should be rejected since there was no significant variation in the response directions or response latencies to no, mild and moderate fluorosis. Nevertheless, severe fluorosis did appear to be judged as significantly less favourable than no, mild, and moderate fluorosis, both in terms of response direction and response latency.

That mild and moderate fluorosis were no different to normal enamel may seem somewhat surprising in the light of past research (McKnight et al., 1998; McKnight et al., 1999; Riordan, 1993a), especially in the case of moderate fluorosis. However, as was discussed in Section 2.5, the reason for the discrepancy between the current investigation and previous work could be the use of full-face photographs in this study rather than the intra-oral photographs or patients typically used. The first study in this thesis, which used standardised extra-oral images of fluorosis, found participants did not spontaneously use or endorse different characteristics to describe normal enamel and mild or moderate fluorosis (Section 2.4). The results of this study not only support this finding, but go even further, since they suggest that participants' attitudes towards normal enamel, mild and moderate fluorosis do not differ in strength or direction, whereas the previous study only considered the direction of their opinion.

The pattern of response latencies (Section 3.3.5) suggest that when a person is of a relatively normal appearance judgements about that person tend to be strong when making a positive attribution and weaker when making a negative attribution. However, if that person has severe fluorosis then this trend is reversed. This implies that we may conform to social norms of politeness, when making social judgements about others unless a person has an appearance that is obviously worse than usual (such as severe fluorosis), in which case we are far quicker to say something negative about that person. This effect may have occurred because severe fluorosis was thought to be a consequence of the persons own neglect, which would support work done by Riordan (1993a) who found that lay people do make this misattribution (about mild and moderate fluorosis although the current study only found this effect for severe fluorosis).

The logistic regressions conducted on individual descriptive themes showed that, in terms of direction of judgement, characteristics related to how physically attractive, careful, clean, happy, healthy, intelligent, kind, reliable, and sociable the images appeared were influenced by level of fluorosis, such that severe fluorosis led to less favourable judgements (although the main effect of level of fluorosis was not significant for the theme of happiness the interaction between level of fluorosis and valence of characteristic for this theme showed that severe fluorosis was judged less favourably).

These findings suggest that severe fluorosis may be attributed to dental neglect, since people with severe fluorosis were perceived as being less clean/more dirty and less careful/more careless, and were possibly judged less reliable/more unreliable on the basis that if they cannot keep their teeth clean, how can they be relied upon for anything else? They may also have been judged as lacking in intelligence for allowing their teeth to become stained. However, the fact people with severe fluorosis were also perceived as being less healthy/more unhealthy suggests that participants may also have attributed severe fluorosis to an illness, which is presumably beyond their control.

They (people with severe fluorosis) were also regarded as being less attractive (or more unattractive), suggesting severe fluorosis was viewed as somewhat disfiguring, and as less sociable (or more unsociable) which could be because the participants assumed that, based on their unusually negative dental appearance, they were not concerned about what others thought of them, or could be because the participants assumed that although they may enjoy socialising, other people would be put off by their appearance. That severe fluorosis led to judgements of the affected individual as being less kind/more unkind could be the result of them being perceived as uncaring of what other people think, and could also be the result of global negative attributions caused by participants' dislike of severely stained teeth. The significant interaction between level of fluorosis and valence of characteristic showed that severe fluorosis led to judgements of the affected individual as less happy/more unhappy, presumably because participants believed their stained teeth would make them unhappy with their appearance.

The linear regressions conducted on individual descriptive themes identified carefulness, happiness, health, and reliability as being particularly influenced, in terms of speed of decision, by level of fluorosis, with severe fluorosis being judged less favourably. This suggests that participants' attitudes were stronger when making attributions about theses themes than the other themes. Therefore assumptions that severe fluorosis was related to dental health, or perhaps to a lack of care with dental hygiene (i.e. participants were not careful with their dental hygiene or were thought to be unreliable because they did not take care of the dental appearance) were particularly strong. Moreover, whatever participants attributed severe fluorosis to, they had strong opinions that it would make the affected individual less happy.

The second main finding of this study was that cueing participants by asking them to pay particular attention to the mouth decreased the likelihood of a "yes" response to a positive characteristic for severe fluorosis and also increased the time taken to make social judgements about somebody with severe fluorosis. However, it did not significantly affect the strength or direction of judgements about mild or moderate fluorosis. This suggests that when participants were cued they may have felt more comfortable saying that a face did not appear to possess a positive characteristic, possibly because being cued led them to assume that different responses were expected for images with unusual teeth, whereas uncued participants were more likely to stick to social norms of politeness and say that the face still appeared to possess the positive characteristic despite having severe fluorosis, even though they may not have believed it. The fact that cued participants appeared to take a longer time to respond to severe fluorosis could be because they assumed severe fluorosis was an experimental manipulation and were attempting to work out the reason(s) for the unusual dental appearance before making their response, whereas uncued participants responded more quickly because they did not deliberate over the cause of the dental appearance, they simply decided that they did not like it. These findings support those of the first study, in which participants made attributions about standardised, extra-oral photographs. The results of that study showed that cueing participants led to a greater number of characteristics varying across levels of fluorosis than, simply asking participants to rate the photographs without cueing them.

Perhaps surprisingly cueing did not have a large impact on the number of participants who reported being aware of mild and moderate fluorosis; with 6 out of 40 uncued participants reporting noticing it compared to 7 out of 39 cued participants, which suggests that mild and moderate fluorosis are not salient features of dental appearance in an extra-oral photograph. The fact that such a small number of participants reported being aware that the dental appearance was varying between normal enamel, and mild and moderate fluorosis suggests that the reason the social judgements and response latencies were so similar for these dental conditions is that the participants did not notice mild and moderate fluorosis. However, all of the participants reported noticing severe fluorosis, and this was reflected in their responses and response latencies. The reason that participants did not notice mild and moderate fluorosis could be that, in the context of an extra-oral photograph, these conditions are simply not noticeable to most people. As was discussed in Section 2.5, studies that have found mild and moderate fluorosis to be a cause for concern have tended to use intra-oral images or patients to display fluorosis. Intra-oral images may exaggerate the effect of fluorosis by removing the context of the face and magnifying the size of the teeth. Using patients could also increase the effects of fluorosis by focusing participants' attention on the teeth. Furthermore, the presence of live subjects may lead to a confounding influence from variables unrelated to fluorosis.

Another possible explanation for participants not reporting being aware of mild and moderate fluorosis is that the current study used filler faces to discourage participants from searching for variations in the target faces. This was intended to make the study more like a real-life meeting, as there is no reason to believe that in a typical social interaction people scrutinize others for minor deviations from normal appearance. Asking participants to alternate between different faces may have prevented them from detecting minor alterations in tooth appearance, which they might have noticed had they only been viewing target faces.

Trials involving severe fluorosis were only shown in the final experimental block of the study. This was because severe fluorosis was obviously different to the other levels of fluorosis (Chapter 2) and would, in all probability, have cued participants to look explicitly at the mouth of the images. Furthermore, it was felt that since the target trials with severe fluorosis were so distinct that when participants began viewing them they might believe that they no longer needed to concentrate on the study and might begin making routine decisions without giving the severe fluorosis trials the same attention they had paid to the other target trials. It was for this reason that half of the filler trials in the final block were replaced with target images with either no, mild or moderate fluorosis.

However, only showing severe fluorosis in the final block of trials may have resulted in a ordering effect when responding to these trials (for example, participants may have felt tired, which may explain why participant's "no" responses to negative characteristic took longer for severe fluorosis than for the other levels, when the converse was expected). Tiredness may also explain the response latencies that were expected. For instance, it was predicted, and found, that when viewing severe fluorosis participants would show longer response latencies when responding "yes" to positive characteristics and "no" to negative characteristics, in comparison to the other levels of fluorosis. This was expected because it would indicate that participants had less positive opinions of a face with severe fluorosis than of the same face with normal enamel or mild or moderate fluorosis. However, the longer response latencies could also have been caused by participants feeling tired, having already completed three experimental blocks. Another possibility for the longer response latencies to severe fluorosis was that participants felt confused about the sudden change in dental appearance (severe fluorosis only appeared in the final block, and most participants were unaware of the differing dental appearances until this point), and consequently hesitated before responding. Therefore, it would be beneficial to run another study with the severe fluorosis trials distributed throughout the four blocks, in order to rule out an ordering effect as the cause of these findings.

When considering the results of the current study there is little scope to relate them to previous work that utilised response latency to measure attitudes because there is very little precedent in the literature of using response latency to measure the strength of social judgements, and none that measured the strength of social judgements of fluorosis. The majority of response latency work has been conducted within the paradigm of the Affective Priming Task (APT) or Implicit Association Test (IAT), which are implicit measures of attitude that ask participants to perform simple categorization tasks rather than to consider their own attitudes (Section 1.4). However, the use of response latency in the current study was not an implicit measure in the usual sense because participants were asked to consider their own attitudes. Although they were unaware that their response latencies were being recorded they were asked to make explicit social judgements. In other words they were aware there attitudes were being measured in terms of their judgments but not in terms of their response latency. In a typical implicit measure they would be unaware their attitudes were being measured at all.

There are a small number of studies that have used response latency to measure social judgements. For example, Dovido, Evans, & Tyler (1986) used response latency to measure social judgements of different racial groups, but, as in the APT and IAT their design only required very simple responses from their participants, and did not require them to consciously deliberate over their own attitudes. Furthermore, their work could not be reproduced in a later study (Judd, Park, Ryan, Brauer, & Kraus, 1995). Another study that successfully measured attitudes to more complex judgements using response latency was conducted by Bassili (1996). He found that participants' response latencies to various questions about social and political issues predicted the stability and pliability (how easily they were persuaded to change their mind) of their opinions. Although, due to the differences in the aims and design of his study and the present study, it is difficult to compare the two.

As was discussed above, in order to validate the current methodology it would be desirable to re-run the study with the severe fluorosis trials dispersed over the four experimental blocks, in order to rule out an ordering effect as the cause of the difference between severe fluorosis and the other levels of fluorosis. It would also be desirable to include dental caries in a study that utilised the current methodology in light of the debate concerning the use of fluoride as a caries preventative agent. This debate weighs the cost of increased exposure to fluoride (increased prevalence of fluorosis) against the benefit of increased exposure (reduced caries – Section 1.3). If it were shown that dental caries are an aesthetic problem equal to or greater than fluorosis then this would help inform debate on the merits of fluoride use.

3.5. Conclusion

Taken together these findings provide further evidence that social judgements about individuals with mild or moderate fluorosis are not different to social judgements about the same individual with normal enamel; but that severe fluorosis does have significant negative impact on social judgements over a range of characteristics. This is true both in terms of the direction of the judgement (whether a person is or is not thought to possess a certain characteristic) and in the strength of the judgement as measured by the length of time taken to respond. Cueing participants to look at the mouth appears to decrease the likelihood of a using a positive characteristic/ increase the likelihood of using a negative characteristic to describe an individual with severe fluorosis and to increase the time taken to make social judgements about individuals with severe fluorosis. Participants did not usually report explicitly noticing mild or moderate fluorosis, even when cued, but all participants noticed severe fluorosis regardless of cueing. A further study that sought to validate the use of the current methodology and to compare attitudes towards mild fluorosis and dental caries would be beneficial.

Chapter 4. The Further Use of the Response Latency Measure to Assess Attitudes to Fluorosis and Dental Caries

4.1. Introduction to the implicit measure of attitudes to fluorosis and dental caries

This, the third and final study conducted in the course of this thesis, was designed to compare social judgements of fluorosis and dental caries. This is of relevance to the debate on the role of fluoride in the prevention of dental caries.

A second subsidiary aim was to determine if the findings of the previous study were the result an ordering effect caused by only including trials involving severe fluorosis in the final experimental block.

The hypothesis of this study was; participants' will be less favourable in their response directions and response latencies when rating dental caries than when rating normal enamel. This is based on previous work that investigated perceptions of dental caries and found them to have a negative impact on participant ratings (Eli et al., 2001; Feng et al., 2001; Newton et al., 2003).

4.2. Method

4.2.1. Recruitment of participants

A similar methodology to that used in the previous study (Section 3.2.8) was used to recruit participants. The administrative office of the Department of Biomedical Sciences e-mailed students on behalf of the experimenters requesting volunteers for the study (Appendix D 1). A different cohort of students to those used in the previous study served as the sample frame. E-mails were sent to first year medical students and to all third year biosciences students (in total 484 e-mails were sent). Additionally posters were displayed in the Department of Biomedical Sciences and the Student Union Building requesting volunteers for the study. Arrangements were made, with students who wished to participate, to meet at mutually convenient time in a room at the Student Union Building, Park Place, Cardiff. All students were kept blind to the fact that the study was connected with dental appearance during the recruitment stage. It was for this reason a venue outside the dental school was selected.

4.2.2. Materials, procedure, and analysis

The procedure, method, materials, procedure, and analysis of this study were identical to those of the second study (Section 3.2) except for five minor differences. These were:

4.2.2.1. The inclusion of dental caries images

The images showing moderate fluorosis were replaced with images showing dental caries. The images of caries were developed in the same way as those described in Sections 2.3.1 and 3.2.1 and are shown in Appendices B 5 and B 10. The first of these images showed an extra-oral photograph of the target male with a non-cavitated lesion on the distal aspect of the right central incisor, and cavitated lesions on the mesial aspects of the upper right and upper left central incisors, and on the buccal aspect of the upper left lateral incisor. The second of these images showed an extra-oral photograph of the target female with cavitated interproximal caries lesions on the mesial aspect of the upper right central incisor, and mesial and distal aspects of the upper left central incisor, and a white spot lesion cervically on the buccal aspect of the upper right lateral incisor.

To comply with the requirements of the programme developed in the previous study (Section 3.2), it was necessary to remove two of the images used in that study (the male and female image at one level of fluorosis) before the images of dental caries could be added. The removal of the moderate fluorosis images was thought most appropriate as normal enamel images served as the reference group and mild fluorosis was deemed as most relevant to the clinical debate on the side effects of fluoride as a caries preventive agent. Retention of the severe fluorosis images would enable comparison with the previous study, and, due to a change in the design of the study (Section 4.2.2.2), would allow investigation into the effects of restricting severe fluorosis trials to the final experimental block in the previous study (Chapter 3).

In the previous study (Chapter 3) the results were described with reference to *level of fluorosis*. Since dental caries is not a level of fluorosis, this study describes the results in terms of *dental appearance* rather than *level of fluorosis*. Similarly, the previous study sometimes used the term *no fluorosis* to denote an image with normal enamel (i.e. an unaltered image). To avoid confusion this study will only refer to *normal enamel* rather than *no fluorosis* when describing images with an unaltered dental appearance since the images of dental caries also depict *no fluorosis*.

4.2.2.2. The order in which the images were displayed

Severe fluorosis was no longer displayed only in the final experimental block. All of the images were placed into random order, and then grouped into four blocks of 108. Therefore it was no longer necessary to replace some of the filler faces in the final block with target faces showing no, mild, or moderate fluorosis. However, as in the previous study the order in which the images were presented within each block was randomised on a participant by participant basis by the SuperLab software. The number of each type of image in each block is shown in Table 4.1.

Table 4.1. Number of each type of image in each block

Block	Normal enamel male	Normal enamel female	Mild fluorosis male	Mild fluorosis female	Dental caries male	Dental caries female	Severe fluorosis male	Severe fluorosis female	Filler faces
1	3	5	6	4	4	5	2	9	70
2	1	3	2	5	5	4	5	3	80
3	5	5	4	4	7	5	6	0	72
4	9	5	6	5	2	4	5	6	66

4.2.2.3. Level of cueing

All participants were cued to the mouth since it was felt that participants' attention would be drawn to the mouth from the first time they viewed severe fluorosis (all participants reported noticing severe fluorosis in the previous study, Section 3.3.6). As the trials appeared in random order within each block the first trial involving severe fluorosis was not presented at the same point for each participant. Therefore, to standardise when participants received cueing to the mouth, they were all cued before beginning the study.

4.2.2.4. Administration of the question sheet

The question sheet used to determine if participants had noticed anything unusual about the images was not administered after the third block. This was carried out in the previous study to investigate if participants noticed changes in the tooth enamel before severe fluorosis appeared; then the question sheet was administered for a second time after the fourth block to investigate if participants noticed changes in the tooth enamel after severe fluorosis appeared. Since all of the images used in the current study were included from the start of the experiment there would have been no point in administering the question sheet twice, therefore it was only administered after completion of the fourth block.

4.2.2.5. Additional question sheet

An extra question sheet was administered at the end of block four (Appendix D 10). This was in addition to the question sheet administered in the second study, and it was designed to assess participants' initial reactions to seeing severe fluorosis and to investigate which of the variations of tooth discolourment participants reported being aware of. Initial reactions to severe fluorosis were explored by asking participants if they felt surprised or confused by the image of severe fluorosis, and also asking them what they believed caused the severe fluorosis. This was to further investigate why, in the previous study, trials involving severe fluorosis tended to produce longer response latencies than trials involving the other levels of fluorosis.

In addition to the questions written on the sheet, two further questions were administered verbally if participants reported being aware of different dental appearances. The first question asked; "How many different types of dental appearance did you notice?", and the second; "Could you please describe the different dental conditions you noticed?". These questions were administered verbally because they depended on the participant reporting being aware of the dental appearance of the images (if they were not aware of it the questions could lead them to mention it). Furthermore, the second question required some interaction between the experimenter and participant since the participant could not be expected to name the dental conditions they had seen, so the experimenter had be sure he fully understood which of the dental appearances the participant was describing. The experimenter wrote down the dental appearances that the participants reported being aware of and this was analysed with the question sheets.

4.2.3. Ethical approval

This study was approved by the Local Research Ethics Committee.

4.3. Results

4.3.1. Participants

Forty students enrolled at Cardiff University participated in this study. In total 15 males and 25 females, with a mean age of 20.93 (± 2.82) years old and an age range of 18-34 years old, participated.

4.3.2. Preparation of data

4.3.2.1. The discarding of inappropriate responses

As in the previous study inappropriate responses were discarded from the data set according to the guidelines described in Section 3.3.2.1. This led to approximately 0.2% of the entire data (18080 responses) set being discarded. None of the participants showed an unusually high rate of erroneous responses, and therefore none of the participants' data were omitted from the analysis. The practice trials and filler trials were also filtered from the data set before it was checked for normality.

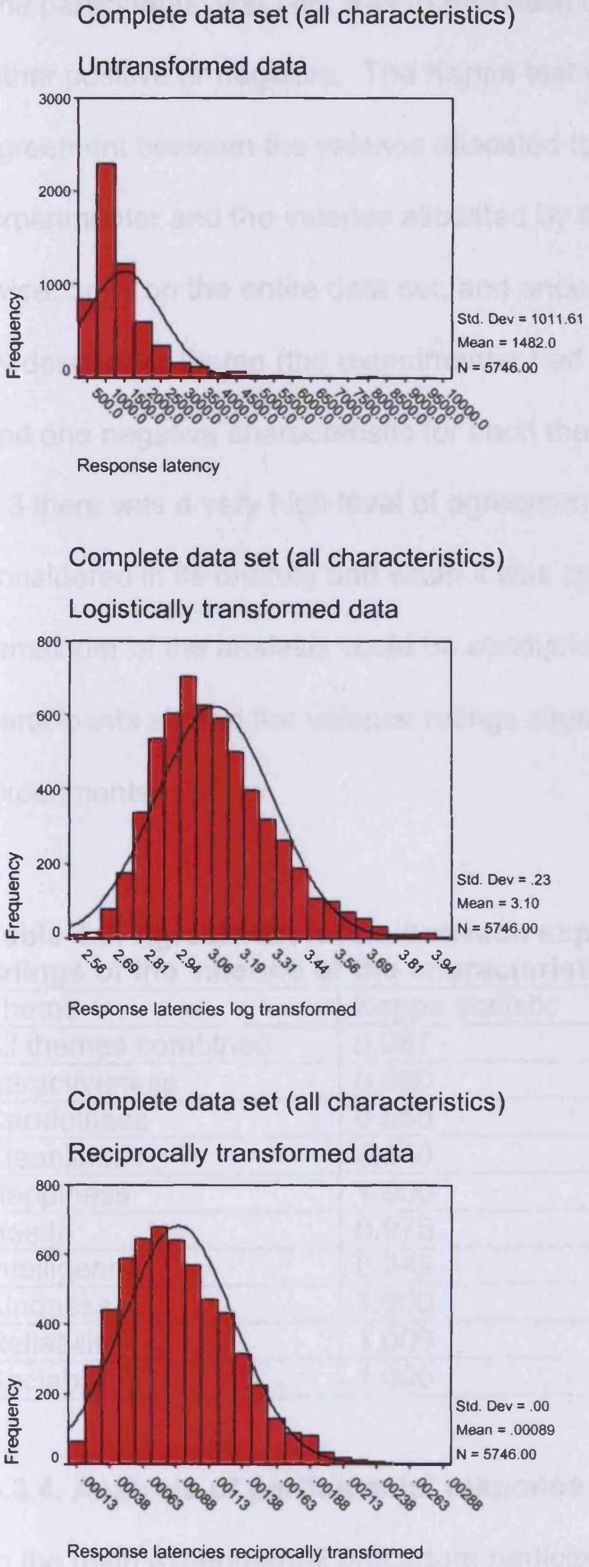
4.3.2.2. The distribution of the response latency data

As in the second study, the data from this study was negatively skewed. Two types of transformation were conducted (logistic and reciprocal), and, as can be seen from Table 4.2 and the graphs in Figure 4.1, logistic transformation was the more appropriate transformation. Therefore, as in the second study, the response latency data were logistically (log_{10}) transformed before analysis.

Table 4.2. Tests of normality of distribution after logistic and reciprocal transformation

Characteristic	Skewness - Transformation method closest to 0 (i.e. closest to normal distribution) shown in **bold**		Kurtosis - Transformation method closest to 0 (i.e. closest to normal distribution) shown in **bold**		Which Graph looked more normally distributed
	Logistic	Reciprocal	Logistic	Reciprocal	
All characteristics	0.683	**0.655**	**0.439**	0.465	Logistic
Attractive	0.939	**0.436**	1.391	**0.238**	No noticeable difference
Careful	**0.710**	0.749	**0.399**	0.777	Logistic
Careless	**0.649**	0.944	**0.546**	1.746	Logistic
Clean	0.713	**0.621**	0.591	**0.541**	No noticeable difference
Dirty	0.899	**0.471**	1.316	**-0.020**	No noticeable difference
Happy	0.729	**0.534**	0.552	**0.204**	No noticeable difference
Healthy	0.875	**0.360**	0.979	**-0.184**	Reciprocal
Intelligent	0.635	**0.527**	0.207	**-0.100**	No noticeable difference
Kind	0.730	**0.519**	**0.205**	0.208	Logistic
Reliable	**0.555**	0.971	**0.163**	1.790	Logistic
Social	**0.715**	0.787	**0.553**	1.105	Logistic
Unattractive	**0.532**	0.816	**0.369**	0.636	Logistic
Unhappy	0.673	**0.434**	**0.050**	-0.420	No noticeable difference
Unhealthy	**0.611**	0.884	**0.461**	1.316	Logistic
Unintelligent	0.602	**0.445**	**-0.031**	-0.441	No noticeable difference
Unkind	**0.634**	0.805	**0.523**	0.866	Logistic
Unreliable	**0.461**	0.752	0.326	**0.199**	No noticeable difference
Unsociable	0.823	**0.469**	0.634	**-0.059**	No noticeable difference

Figure 4.1. Histograms of the data distribution for the entire data set before transformation and after logistic and reciprocal transformation



4.3.3. Participants valence ratings of the characteristics

The participants' first task was to rate each of the characteristics as being either positive or negative. The Kappa test was used to determine the level of agreement between the valence allocated to each characteristic by the experimenter and the valence allocated by the participants. The test was run twice, once on the entire data set, and once after the data set had been split by descriptive theme (the experimenter had deliberately selected one positive and one negative characteristic for each theme). As can be seen from Table 4.3 there was a very high level of agreement, both when the data set was considered in its entirety and when it was split by theme. Therefore the remainder of the analysis could be conducted in the knowledge that participants shared the valence ratings allocated to each characteristic by the experimenter.

Table 4.3. Agreement levels between experimenter and participant ratings of the valence of the characteristics

Theme	Kappa statistic	Significance
All themes combined	0.967	p<0.001
Attractiveness	0.950	p<0.001
Carefulness	0.850	p<0.001
Cleanliness	0.950	p<0.001
Happiness	1.000	p<0.001
Health	0.975	p<0.001
Intelligence	0.949	p<0.001
Kindness	1.000	p<0.001
Reliability	1.000	p<0.001
Sociability	1.000	p<0.001

4.3.4. Analysis of participants' response direction

In the main experimental procedure participants pressed a "yes" or "no" key to show whether they thought a characteristic described a face. The percentage of "yes" responses to each characteristic, for each dental appearance are shown in Table 4.4.

Table 4.4. Percentage of "yes" responses [a] to each characteristic by dental appearance

Characteristic	N [b]	Normal enamel	Mild fluorosis	Severe fluorosis	Dental caries
Attractive	320	61.3	63.8	11.3	31.3
Unattractive	320	21.3	35.0	71.3	63.8
Careful	320	78.8	76.3	23.8	37.5
Careless	318	21.5	19.0	55.0	51.3
Clean	319	83.8	83.8	18.8	32.9
Dirty	319	13.8	8.8	72.5	59.5
Happy	320	90.0	87.5	68.8	76.3
Unhappy	319	7.6	3.8	23.8	12.5
Healthy	319	87.5	86.3	26.3	39.2
Unhealthy	320	12.5	12.5	66.3	46.3
Intelligent	320	85.0	72.5	43.8	51.3
Unintelligent	318	11.3	14.1	36.3	21.3
Kind	320	92.5	88.8	75.0	77.5
Unkind	318	6.3	7.5	9.0	8.8
Reliable	319	87.5	78.8	39.2	58.8
Unreliable	317	8.8	10.3	41.8	23.8
Social	320	86.3	87.5	52.5	67.5
Unsocial	320	1.3	7.5	35.0	17.5

[a] The remaining percentage are "no" responses.
[b] N varies slightly because responses considered erroneous were omitted.

Participants' response direction was analysed using logistic regression where the dependent variable was the likelihood of a "yes" response (the logistic regression equation is shown in Appendix F 1). Table 4.5 summarises the logistic regression conducted on the entire data set. Since all participants were cued in this study there is no main effect of cueing, or interactions involving cueing.

Table 4.5. Two-level logistic regression model of response direction for all characteristics combined

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	79.76	(39.62, 160.57)	$p<0.001$
Dental appearance			$p<0.001$
Normal enamel	1[a]	R	
Mild fluorosis	1.17	(0.84, 1.65)	
Dental caries	4.92	(3.63, 6.68)	
Severe fluorosis	9.03	(6.66, 12.23)	
Valence of Characteristic* Dental appearance			$p<0.001$
Normal enamel	1[a]	R	
Mild fluorosis	0.66	(0.42, 1.05)	
Dental caries	0.03	(0.02, 0.05)	
Severe fluorosis	0.01	(0.01, 0.01)	

[a] Normal enamel is the reference category

Table 4.5 shows that the main effects of valence of characteristic and dental appearance were both significant, as was the interaction between them. Participants were more likely to give a "yes" response to a positive characteristic than to a negative characteristic, and were more likely to give a "yes" response to severe fluorosis or caries than to normal enamel, with "yes" responses most common for severe fluorosis.

The interaction between dental appearance and valence of characteristic shows that participants were significantly less likely to give a "yes" response to a positive characteristic for severe fluorosis or caries, with severe fluorosis showing the lowest likelihood of a "yes" response to a positive characteristic. The interaction between dental appearance and valence of characteristic showed that mild fluorosis did not appear to be different to normal enamel. This suggests that although "yes" responses were more common for severe fluorosis and dental caries than for normal enamel this can be largely attributed to the fact there were far more "yes" responses to trials involving negative characteristics when judging severe fluorosis and caries than when judging trials involving normal enamel.

Table 4.6 is a prediction table created using the logistic regression conducted on the entire data set. This table clearly illustrates that, mild fluorosis had a very similar percentage of "yes" responses to normal enamel, and that both had more "yes" responses for positive characteristics than negative characteristics. However, in comparison with normal enamel, "yes" responses for severe fluorosis were far less likely for positive characteristics, but far more likely for negative characteristics, such that there were more "yes" responses for negative characteristics than positive characteristics. This was also true of caries, but to a lesser extent, so that caries appeared to be less favourable than normal enamel or mild fluorosis but more favourable than severe fluorosis.

Table 4.6. Prediction table for likelihood of a "yes" response to positive and negative characteristics by dental appearance

Dental appearance	Percentage likelihood of a "yes" response	
	Positive characteristics	Negative characteristics
Normal enamel	88	8
Mild fluorosis	85	10
Severe fluorosis	37	45
Dental caries	52	30

Logistic regressions were conducted on the data after it was split into individual descriptive themes. The p-values of the main effects and interactions in these regressions are summarised in Table 4.7 (MLwiN could not converge two of the logistic regressions after the interaction between valence of characteristic and dental appearance had been added, probably due to small numbers in some of the cells, in these instances only the main effects are reported). Although there was some variation in the significance levels, the direction of the responses were the same as those described in the regression conducted on the entire data set, i.e. caries and severe fluorosis appeared to be judged less favourably than normal enamel, but mild fluorosis was not judged differently to normal enamel. The logistic regression tables for the individual descriptive themes are in Appendix F 5.

Table 4.7. Summary of p-values for each variable in the individual logistic regressions [a]

Regression analysis	Valence of characteristic	Dental appearance	Valence of characteristic* Dental appearance
All themes combined	p<0.001	p<0.001	p<0.001
Attractiveness	p<0.001	p<0.001	p<0.001
Carefulness	p<0.001	p<0.001	p<0.001
Cleanliness	p<0.001	p<0.001	p<0.001
Happiness	p<0.001	p<0.001	p<0.001
Health	p<0.001	ns	-
Intelligence	p<0.001	p<0.001	p<0.001
Kindness	p<0.001	p<0.05	-
Reliability	p<0.001	p<0.001	p<0.001
Sociability	p<0.001	p<0.001	p<0.001

[a] Some of the cells are empty because not all of the variables were fitted in every regression, see Section 3.2.12.1.

Participants' social judgements on the themes of physical attractiveness, carefulness, cleanliness, happiness, intelligence, kindness, reliability, and sociability all showed the trend for severe fluorosis and caries to be judged less favourably than normal enamel, such that, for example, participants would be less likely to say that a person was attractive if they had severe fluorosis or caries (compared to normal enamel) and more likely to say that they were unattractive. Mild fluorosis was not shown to be different to normal enamel in these judgements. The only judgements that did not reveal significant differences in the interaction between dental appearance and valence of characteristic were those judgements related to how healthy and kind a person appeared to be.

These findings suggest that the social judgements made about an individual are not altered if that individual has mild fluorosis, but do change unfavourably if they have severe fluorosis or dental caries; and that this is true of social judgements made about a range of characteristics. Furthermore, severe fluorosis appears to be judged less favourably then dental caries.

4.3.5. Analysis of participants' response latencies

In addition to participants response direction (whether they pressed the "yes" or "no" key) their response latency (the time between the presentation of the characteristic and the participant pressing a key) was recorded. This was subsequently analysed separately from response direction. This section (4.3.5) describes the analysis of participants' response latency.

Descriptive statistics of the raw response latencies for the entire data set are shown in Table 4.8. Tables showing the descriptive statistics of the raw response latencies for responses to each individual characteristic are shown in Appendices F 3 and F 4.

Table 4.8. Response latencies (milliseconds), by dental appearance for all characteristics combined

Valence of characteristic	Response direction	Normal enamel			Mild fluorosis			Severe fluorosis			Dental caries		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Positive	Yes	1282	1001	341-7631	1304	1012	350-6259	1657	1342	490-9884	1611	1291	501-9904
	No	1640	1387	521-6570	1633	1292	521-7361	1390	1132	520-5848	1544	1182	491-9244
Negative	Yes	2081	1622	350-6710	1870	1583	561-6769	1604	1282	531-8682	1512	1241	420-5929
	No	1359	1102	350-7241	1394	1132	451-7561	1697	1312	430-8823	1521	1202	490-8232

The purpose of recording response latency was to use it as a measure of the strength of participants' attitudes (with shorter response latencies indicating greater attitude strength). The response latency data was analysed using linear regression where the dependent variable was the transformed (to the $\log_{10}$) response latency. The final model selected for the linear regression is shown in Appendix F 2. The regression table for this equation is Table 4.9. Since all participants were cued in this study there is no main effect of cueing, or interactions involving cueing. The impact of the main effects and interactions of interest are summarised in Table 4.10.

Table 4.9. Two-level linear regression model showing the differences in response latency for all characteristics combined

	Coefficient	95% Confidence Interval	Significance
Dental appearance			p<0.001
Normal enamel	0 [a]	R	
Mild fluorosis	0.007	(-0.014, 0.028)	
Dental caries	0.036	(0.014, 0.060)	
Severe fluorosis	0.073	(0.049, 0.097)	
Response direction	0.137	(0.092, 0.181)	p<0.001
Valence of characteristic	0.058	(0.020, 0.096)	p<0.01
Valence of characteristic* Response direction	-0.233	(-0.292, -0.175)	p<0.001
Dental appearance* Valence of characteristic			ns
Normal enamel	0 [a]	R	
Mild fluorosis	-0.008	(-0.060, 0.043)	
Dental caries	-0.048	(-0.095, -0.002)	
Severe fluorosis	-0.111	(-0.016, -0.065)	
Dental appearance *Response direction			p<0.01
Normal enamel	0 [a]	R	
Mild fluorosis	-0.024	(-0.085, 0.036)	
Dental caries	-0.128	(-0.181, -0.075)	
Severe fluorosis	-0.143	(-0.195, -0.090)	
Dental appearance* Valence of characteristic *Response direction			p<0.001
Normal enamel	0 [a]	R	
Mild fluorosis	0.038	(-0.042, 0.118)	
Dental caries	0.229	(0.158, 0.300)	
Severe fluorosis	0.277	(0.206, 0.348)	

[a] Normal enamel is the reference category

As can be seen from Table 4.9 dental appearance had a significant effect on response latency, and, based on the coefficients and confidence intervals, it would appear that this difference is attributable to judgements involving severe fluorosis and caries being slower than judgements involving normal enamel, with judgements involving severe fluorosis showing the slowest response latencies. Mild fluorosis does not appear to have significantly different response latencies to normal enamel. The precise nature of the difference is illuminated in Table 4.10.

Table 4.10. Change in logged response latency (milliseconds) associated with dental appearance, valence of characteristic and response direction

Response direction	Dental appearance by valence of characteristic							
	Normal enamel		Mild fluorosis		Severe fluorosis		Dental caries	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
No	0	0.058	0.007	0.057	0.073	0.020	0.036	0.046
Yes	0.137	-0.039	0.119	-0.020	0.067	0.131	0.045	0.087

Table 4.10 shows the impact on logged response latency of response direction and valence of characteristic across the different dental appearances. The basic model of the linear regression was designed to show response latencies for "no" responses to negative characteristics, for normal enamel, therefore the value in this cell is 0. When judging faces with normal enamel or with mild fluorosis the fastest responses (i.e. the smallest numbers) were for "yes" responses to positive characteristics. "No" responses to negative characteristics were slightly slower, and then there was a larger increase in the time taken for "no" responses to positive characteristics, and a slower response still for "yes" responses to negative characteristics. In other words people were quickest to say something positive about a person, and were almost as quick to say that something negative was not true of a person, indicating their attitudes were strongest for these responses. They took a longer time when saying that a person did not appear to possess a positive attribute (indicating a weaker attitude), and their slowest responses (and therefore, weakest attitudes) were when endorsing something negative about a person. This pattern suggests there were social norms of politeness evident which meant participants were faster to say something complimentary about the images than to say something critical. This effect was slightly stronger for normal enamel than for mild fluorosis, but the confidence intervals indicate it was not significantly stronger.

There was a different pattern in the response latencies to dental caries and severe fluorosis. The fastest responses to severe fluorosis were for "no" responses to positive characteristics, followed by "yes" responses to negative characteristics. There was a slight increase in response latency for "no" responses to negative characteristics, and finally the slowest response latencies to severe fluorosis were for "yes" responses to positive characteristics. This indicates that, when judging severe fluorosis, participants strongest attitudes were when withholding a positive statement (i.e. they were quickest to say "no" to a positive characteristic) and were also strong when saying something negative was true of a person with severe fluorosis. Whereas, their longest response latency (reflecting their weakest opinion) was to say something positive was true of a person with severe fluorosis. This pattern of results suggests that while participants were faster to make complimentary judgements than critical judgements about normal enamel and mild fluorosis (which indicates stronger attitudes for complimentary judgements); the opposite was true for judgements about severe fluorosis.

Table 4.10 shows that, when judging dental caries, participants' response latencies were shortest (indicating their strongest attitudes) for "no" responses to negative characteristics (which is similar to the social norms of politeness evident when rating normal enamel or mild fluorosis), and their longest response latencies (weakest attitudes) were given for "yes" responses to positive characteristics (which is similar to the pattern of response latencies for severe fluorosis). Response latencies to "yes" responses to negative characteristics and "no" responses to positive characteristics were almost identical and fall between the two extremes. This pattern of response latencies is similar to the pattern in the responses directions (see Table 4.6) in that it appears to indicate caries is viewed less favourably than normal enamel or mild fluorosis, but more favourably than severe fluorosis.

Further linear regressions were conducted on the data after it was split into individual descriptive themes. As in the model fitted to the entire data set, variables were added in groups to see if they explained significantly more of the variance. Thus not all of the descriptive themes have p-values for each variable (Section 3.2.12.2 and Table 3.5). The linear regression tables for the individual descriptive themes are in Appendix F 6. The p-values of the main effects and interactions in these regressions are summarised in Table 4.11.

Table 4.11. Summary of p-values for each variable in the individual linear regressions [a]

Regression analysis	Dental appearance	Response direction	Valence of characteristic	Valence of characteristic* Response direction	Dental appearance *Response direction	Dental appearance* Valence of characteristic	Dental appearance* Valence of characteristic *Response direction
All themes combined	p<0.001	p<0.001	p<0.01	p<0.001	p<0.001	p<0.001	p<0.001
Attractiveness	p<0.05	p<0.001	ns	p<0.001	ns	p<0.05	p<0.01
Carefulness	p<0.05	p<0.05	ns	p<0.05	-	-	-
Cleanliness	p<0.001	p<0.01	ns	p<0.001	p<0.001	p<0.001	p<0.001
Happiness	p<0.001	ns	ns	p<0.05	-	-	-
Health	p<0.05	ns	ns	p<0.05	p<0.05	p<0.01	p<0.001
Intelligence	p<0.01	p<0.05	ns	p<0.05	ns	p<0.01	ns
Kindness	p<0.001	p<0.001	p<0.05	p<0.001	-	-	-
Reliability	p<0.05	ns	ns	ns	-	-	-
Sociability	ns	ns	ns	ns	-	-	-

[a] Some of the cells are empty because not all of the variables were fitted in every regression, see Section 3.2.12.2.

The three-way interaction between dental appearance, response direction, and valence of characteristic described above was replicated in the regressions conducted on the themes of physical attractiveness, cleanliness, and health, suggesting attributions made about these themes were particularly strong. Although there was some variation in the significance levels, the direction of the three-way interaction for each individual theme was the same as that described in the regression conducted on the entire data set.

4.3.6. The question sheet assessment

The results of the question sheets that assessed how aware participants were of the varying dental appearances showed that all 40 participants noticed the normal enamel, dental caries and severe fluorosis, but only 26 participants noticed mild fluorosis. Of the 40 participants 34 reported being surprised the first time they saw severe fluorosis, but only 11 said they were confused. The majority of the participants (28) thought that the severe fluorosis was a deliberate manipulation designed to look like unhealthy teeth the first time they saw it, and a further seven thought that the severe fluorosis was a deliberate manipulation designed to look like staining caused by smoking, food, etc. Two of the participants thought the severe fluorosis image was a genuine photograph the first time they saw it, and two of them thought it was a computer glitch. Only one of the participants did not think the severe fluorosis was due to any of these reasons the first time they saw it; that participant thought they were supposed to be paying close attention to the shape of the mouth and that the severe fluorosis was designed to draw their attention away from this.

These findings suggest severe fluorosis and dental caries are explicitly noticed to a greater degree than mild fluorosis by the sample group. The increase in the number of participants noticing mild fluorosis from the previous study (26 out of 40 in the present study compared to 7 out of 39 in the cued condition of the previous study, and 13 out of 79 in the whole of the previous study) suggests that including severe fluorosis from the beginning of the study focused participants' attention on the teeth to the extent to which they became more likely to notice small deviations from normal enamel (i.e. mild fluorosis). This has two implications, the first and more important is that even when the majority of participants were explicitly aware of mild fluorosis they still did not make different social judgements based on whether the person they were judging had normal enamel or mild fluorosis. The second is that asking participants to pay particular attention to the mouth is a less effective way of cueing them to fluorosis than including images of severe fluorosis.

4.4. Discussion

The present study had two main aims, therefore the results of this study will be discussed with reference to these aims. The first aim was to compare social judgements of fluorosis and dental caries. The results of this study showed that participants judgements of dental caries were less favourable than their judgements of mild fluorosis and normal enamel, but more favourable than severe fluorosis. This is true both in the direction of participants' judgements and the strength of their judgements (as measured by response latency). This finding means that the experimental hypothesis can be accepted since participants' response directions and response latencies both showed dental caries to be rated less favourably than normal enamel.

The finding that dental caries was judged less favourably than mild fluorosis has implications for the debate on the role of fluoride in the prevention of dental caries. The use of fluoride supplements, or the fluoridation of drinking water, may result in an increased prevalence of mild fluorosis but reduced prevalence of dental caries (Section 1.3). Therefore fluorosis can be considered the cost of fluoridation and reduced dental caries the benefit. However, the current results suggest that the supposed cost of fluoridation, mild fluorosis - which is an aesthetic issue, may be unproblematic and even preferable to the aesthetic cost of dental caries, the prevalence of which is likely to be higher in non-fluoridated communities.

When interpreting these results the images used in this study should be considered. Since the current study sought to compare dental caries with mild fluorosis it was necessary to develop standardised images showing fluorosis and dental caries (Appendix B). The images were developed based on genuine clinical photographs, but it is important to be aware that there is some variation in the appearance of both fluorosis and dental caries, and caries in particular presents problems regarding whether it should be showed as treated or untreated (the current study used images of untreated dental caries). It is also necessary to acknowledge that it is not truly possible to compare "equivalent" levels of fluorosis and caries, because they are different conditions. Therefore it was necessary for the experimenters to exercise their judgement in the development of these images.

The second aim of this study was to investigate whether the longer response latencies for severe fluorosis in the previous study were the result of an ordering effect caused by only including severe fluorosis trials in the final experimental block. The study was successful in this aim since it found the longer response latencies do not appear to have been caused by an ordering effect. The longer response latencies for severe fluorosis when participants gave "yes" responses to positive characteristics or "no" responses to negative characteristics study were replicated in this study, in which severe fluorosis trials were dispersed throughout the four experimental blocks.

In fact if anything, the ordering effect in the previous study appears to have diluted the trend rather than caused it, since the previous study showed that some response latencies to severe fluorosis were longer than response latencies to the other dental conditions even when they were expected to be shorter (i.e. "no" responses to positive characteristics were expected to shorter for severe fluorosis than for the other levels, but were longer). However, this unexpected finding did not occur in the current study.

The longer response latencies may, however, be partially accounted for by the effects of surprise or confusion, since most of the participants (34) reported being surprised the first time they saw severe fluorosis and 11 of them reported feeling confused. This may have caused them to pause before responding to trials involving severe fluorosis. However, it is unlikely that this effect would have had a large impact on their overall response latencies since it would probably only have affected the first trial or two involving severe fluorosis and only a small number of participants' responses were over the acceptable response latency limit. Furthermore, most participants believed severe fluorosis was a deliberate manipulation portraying unhealthy teeth (28 participants) or staining (seven participants). It therefore seems more likely that the longer response latencies reflect participants' attitudes towards unhealthy or stained teeth rather than their confusion or surprise, because these assumptions (that teeth were unhealthy or stained) would be relevant for every severe fluorosis trial, whereas confusion or surprise would be unlikely to last beyond the first trial or two involving severe fluorosis.

Having discussed the outcome of the study in relation to the two principal aims, the results of this study will now be compared to previous work with regard to the participants awareness of mild fluorosis, the pattern of participants response latencies, the individual descriptive themes that were particularly relevant to dental appearance, and the use of the descriptive themes with images of dental caries.

4.4.1. Participants awareness of mild fluorosis

It is worth noting that participants' response directions and response latencies to mild fluorosis were similar to those for normal enamel, even though the majority of participants (26 out of 40) reported being aware of whitened enamel in some images (i.e. mild fluorosis). This suggests that even when people are aware of mild fluorosis in others they do not perceive it negatively, which supports work by Riordan (1993a), Ellwood & O'Mullane (1995), and Hawley et al. (1996), but contradicts work by McKnight et al. (1998) and McKnight et al. (1999).

4.4.2. The pattern of participants' response latencies

The exact pattern of response latencies in the interaction between dental appearance, valence of characteristic, and response direction was slightly different in this study than in the last. For example, in this study participants were quickest to give a "yes" response to a positive characteristic and almost as quick to give a "no" response to a negative characteristic when viewing normal enamel or mild fluorosis, whereas in the previous study this was the other way around. However, the social norms of politeness found in the previous study were still evident in this study, i.e. people were quicker to say something positive or to reject a negative statement about a person, than to say something negative or withhold a positive statement about a person unless they had a dental condition that breached social norms (i.e. unless they had severe fluorosis or dental caries). It may be that this trend can be taken as a validation of the methodology since the response latencies do not vary randomly, but rather seem to reflect underlying social norms of politeness.

4.4.3. The individual descriptive themes identified as particularly relevant to dental appearance

The logistic regressions conducted on individual descriptive themes showed that, in terms of direction of judgement, characteristics related to how physically attractive, careful, clean, happy, intelligent, kind, reliable and sociable the target images appeared were influenced by dental appearance, such that caries and severe fluorosis led to less favourable judgements than normal enamel or mild fluorosis, and severe fluorosis was less favourable than caries. But characteristics related to how healthy people appeared to be did not seem to be influenced by dental appearance.

These findings are very similar to those of the previous study except for the surprising finding that dental appearance did not significantly effect judgements about health. Given the results of both of the previous studies and the fact that poor dental health can result in tooth discolouration (e.g. caries), it may have been expected that health would have been one of the themes most effected by dental appearance. Perhaps the reason it was not was that participants assumed that the discolouration was caused by neglect, and was therefore an indication of attitude to dental care rather than health. Other than this, the only other differences between this study and the previous one were that the logistic regression conducted on the theme of happiness showed a significant main effect of dental appearance in this study but not in the last. Additionally, the interaction between dental appearance and valence of characteristic could not be calculated for the theme of kindness in this study (probably due to small numbers in the cells), but it was significant in the previous study.

In the linear regressions of response latencies, conducted on individual descriptive themes, the three way interaction between dental appearance, valence of characteristic and response direction (the nature of which is described in Section 4.3.5) reached significance for the themes of physical attractiveness, cleanliness, and health. That the response latencies to these themes were particularly affected by tooth discolouration may not be a surprise since they are more obviously related to appearance than most of the other themes, which are more related to personality. However, with the exception of health, these themes did not reach significance in the linear regressions conducted in the previous study. The themes in which the three-way interaction between level of fluorosis, valence of characteristic and response direction reached significance in that study were carefulness, happiness, health, and reliability. It should be noted though, that is likely that in both this study and the previous one, the three-way interaction was evident to some extent in the themes that did not reach significance, since when the entire data set was analysed the three way interaction was highly significant ($p<0.001$ in both studies). It is unlikely that just three of the nine themes (or four of the nine themes in the previous study) could have caused such a large finding in the whole data set.

4.4.4. The use of the descriptive themes with images of dental caries

The characteristics used in this study were identified as relevant to fluorosis by the first study of this thesis (Chapter 2) and were generated using images of normal enamel and mild, moderate and severe fluorosis but did not include images of caries. Therefore, it could be possible that the characteristics used in this study do not cover all of the attributions that are made on the basis of dental caries. However, this is unlikely since previous work that has investigated the social consequences of tooth decay has identified characteristics similar to those used in the current study as being important to caries. For instance, Eli et al. (2001) identified perceptions of three categories of traits as being affected by the presence of tooth decay; these were; aesthetic, social, and professional. The aesthetic category would appear to be covered in the present study by the themes of attractiveness, and cleanliness (Eli et al., 2001, included in “clean” and “dirty” in the category of aesthetic). The social category could be covered by the theme of sociability, and the professional category could be covered by the themes of reliability, intelligence, and carefulness.

Two studies by Feng et al. (2001) and Newton et al. (2003) found that caries affected perceptions of adjustment, social competence and intellectual competence. Adjustment could be covered by the theme of happiness in the current study, social competence by the theme of sociability, and intellectual competence by the theme of intelligence. Therefore it appears the current study has covered the themes that are relevant to caries when people are asked to attribute characteristics to others.

4.5. Conclusion

These findings provide further evidence that social judgements about individuals with mild fluorosis are not different to social judgements about the same individual with normal enamel. Furthermore, they suggest that dental caries does have a negative impact on the social judgements made, and that therefore mild fluorosis would be preferable to dental caries. Severe fluorosis appears to have a greater negative impact on social judgements than dental caries. These findings account for both the direction of participants' judgements (whether a person is or is not thought to possess a certain characteristic) and the strength of their judgements as measured by the length of time taken to respond. This study also showed that the differences between normal enamel and severe fluorosis in response latency in the previous study were unlikely to have been caused by an ordering effect.

Chapter 5. General Discussion

The general aims of this thesis were to use both explicit and implicit measures of attitude to identify characteristics attributed to others that varied according to their dental appearance. Over three experiments the research has found that, when judging standardised, extra-oral images, people do not attribute characteristics to others with mild or moderate fluorosis that differ from those they attribute to others with normal enamel. However, when judging an individual with severe fluorosis or caviated dental caries people do make less favourable attributions, with severe fluorosis eliciting less favourable judgements than caries. Furthermore, the strength of people's attributions (as implicitly measured by response latency) do not differ between normal enamel, mild fluorosis and moderate fluorosis, but do show differences between these dental appearances and severe fluorosis or dental caries; such that severe fluorosis and dental caries are associated with weaker positive attitudes and stronger negative attitudes.

Cueing participants, by asking them to pay particular attention to the mouth of an extra-oral image, appears to influence responses to photographs of severe fluorosis, but does not appear to affect judgements about mild or moderate fluorosis. Asking people to make attributions based on intra-oral photographs resulted in mild fluorosis being judged less favourably than normal enamel and moderate fluorosis for some characteristics, including attractive, clean, careful, happy, intelligent, and sociable.

These findings suggest that mild and moderate fluorosis are not regarded as a cause for concern by the public, and that mild fluorosis is preferable to caviated dental caries in terms of dental appearance, which has implications for the debate on the use of fluoride as a caries preventive agent. They also indicate severe fluorosis is perceived as a cause for concern, and that both severe fluorosis and dental caries can lead to the attribution of negative characteristics, or the non-attribution of positive characteristics that go beyond the aesthetic. This research provides evidence that previous work, which found mild and moderate fluorosis to be perceived as problematic (Ellwood & O'Mullane, 1995; Hawley et al., 1996; McKnight et al., 1998; McKnight et al., 1999; Riordan, 1993a) may have been the result of using intra-oral images, or of cueing participants to the teeth of a patient with fluorosis. Doing this can magnify the aesthetic impact of fluorosis beyond that which would be evident if participants judged extra-oral images, which are closer to representing a typical, real-life social interaction.

Based on the results of the three studies described in this thesis, it appears that there are a variety of attributions made about tooth discolouration, and that these go beyond judgements about physical attractiveness and may include assumptions about carefulness, cleanliness, happiness, health, intelligence, kindness, reliability, and sociability. This supports the studies described in the literature review (Section 1.2), which suggest people make attributions about the personal characteristics of others, based on their appearance, which go beyond the aesthetic. For example Eagly et al. (1991) found that perceptions of potency, adjustment, intellectual competence, and social competence were associated with physical attractiveness. The current study, which found that people's attributions of characteristics to others varied according to their dental appearance, relates to work by Eli et al. (2001), Feng et al. (2001), Newton et al. (2003), Shaw, (1981), Shaw & Humphreys (1982), and Shaw et al. (1985) which found dental appearance to be important in the attributions people make about others with decayed teeth or abnormal tooth alignment.

The fact that such a wide variety of characteristics were, or were not, attributed based on dental fluorosis could have many different implications for how people behave towards individuals with fluorosis. For example, dating and friendship choices could be influenced by attributions based on severe fluorosis, as could professional decisions on an individual's competence, if they were thought less intelligent, reliable or careful due to their stained and pitted teeth. However, it should be noted that the attribution of characteristics was only influenced by severe fluorosis, which is very rare in the U.K. (Milsom & Mitropoulos, 1990; Office for National Statistics, 2003).

The principal goal of the research conducted in this thesis was to examine the social impact of dental fluorosis. As was described in Section 1.4.4, it was decided the best way to achieve this aim was to develop an implicit measure of how social judgements made about individuals varied according to their level of fluorosis. Literature related to two of the foremost implicit measures used in attitude measurement, the Affective Priming Task (APT) and the Implicit Association Test (IAT), was discussed in Section 1.4. It was concluded that there was no previous methodological design that was directly applicable to the current aim. Therefore, it was decided to develop a new measure based on one or both of the APT and IAT to meet the needs of this research. After considering both measures it was thought that a technique based on the APT was more suitable for the current research. This design is described in detail in Chapter 3. The rationale behind this decision, and the success of the methodology developed are now discussed.

5.1. The rationale for developing the methodology used in this research

When designing the methodology both the research question and the materials to be used were considered. It was decided that the most effective way of investigating attitudes to fluorosis would be to develop standardised extra-oral photographs of individuals with different levels of fluorosis, and this was undertaken as part of the first study (Chapter 2). Furthermore, characteristics that were relevant to fluorosis were identified by the extent to which they were attributed to standardised photographs with varying levels of fluorosis (Chapter 2). These photographs and characteristics served as materials in the implicit measure of attitudes.

One advantage of using photographs in an implicit measure of attitudes was that images have been found to be more potent stimuli than words in both the Affective Priming Task (Fazio, 1993a) and the Implicit Association Test (Swanson et al., 2001). Neither measure appears to have been used before with an image modified in order to study how variations within the same background image effect participants' judgements, which is what the current study attempted to do. However, the APT has been shown to work successfully when the prime words are presented at subliminal speeds, (Niedenthal, 1990; Wittenbrink et al., 1997) which suggests the APT may be sensitive enough to discern between small variations in the images. Therefore, in terms of the materials it was desirable to use, the APT had a small advantage over the IAT.

5.1.1. Advantages and disadvantages of using the Implicit Association Test (IAT)

Given that the aim of the research was to assess social judgements, it might seem that the IAT was a more suitable model, since it measures associations, and would therefore be measuring the associations between the photographed individuals and the characteristics, which seems a reasonable way to examine social judgements. Whereas if a typical APT were used, it would be measuring the extent to which both the photographed individuals and the characteristics shared the same valence; and saying that the person X and characteristic Y are both positive is not the same thing as saying person X is characteristic Y. Furthermore, the IAT is probably the best known of the implicit measures (Fazio & Olson, 2003, p.298) and since its introduction the IAT has been the measure of choice for the majority of studies that implicitly measured attitudes, despite being more recently introduced than the APT.

However, when the materials to be used in the current research were considered the IAT did have some drawbacks. The IAT asks participants to sort four concepts using two responses. This means that it would only be possible to use two levels of fluorosis and two adjectives per IAT, for instance one IAT could compare normal enamel and mild fluorosis on the descriptive theme of attractiveness. In this IAT participants would be asked to classify the faces as having "no fluorosis" or "mild fluorosis", and to classify words as meaning "attractive" or "unattractive". In order to compare three levels of fluorosis to normal enamel, for all nine descriptive themes, it would require 27 IATs. Clearly this would require either a very large number of participants or would require each participant to spend a very long period of time on the task.

A further problem with using an IAT in the current study is that the photographed individuals would need to be easily classified according to level of fluorosis. Distinguishing between mild and moderate fluorosis and distinguishing either of them from normal enamel may not be easy for the untrained eye. It would also create the further problem of explaining to participants what fluorosis is, and how to recognise one level from another, which would certainly make the study unlike a genuine social interaction, and thus detract from the goal of the study. It might have been possible to ask participants to judge the faces based on some other criteria such as male or female, or previously shown face versus new face and then to look for interactions with level of fluorosis. This was done previously by Castelli et al. (2004, experiment 2), when investigating how participants knowledge of which category a person belonged to influenced their responses when shown that person in an unrelated task. However, this would complicate the IAT and given that it is already unclear to what extent the IAT measures associations between the concepts (Section 1.4.3.2), this is undesirable. Indeed, Castelli et al. (2004) chose to use the APT rather than the IAT in the majority of their experiments because they felt it was better established as a measure of spontaneous affective reactions, presumably this was because there is a greater body of research concerning the mechanisms underlying affective priming than the mechanisms behind the IAT.

Given these issues with the IAT, the APT was considered in more detail before making a final decision.

5.1.2. Advantages and disadvantages of using the Affective Priming Task (APT)

The APT has advantages over the IAT since it does not require participants to be aware of fluorosis and there is a greater understanding of how affective priming works than of how the IAT works (Fazio & Olson, 2003). However, there were three other problems with using the APT in the current research. The first was that it appears to measure the extent to which a photographed target face and a characteristic are both positive or negative, and not the extent to which the face appears to possess the characteristic. The second problem was that there would be two response outcomes to analyse (response direction and response latency). The third and final problem was that in a typical APT there would only be 300ms between the presentation of the face and the presentation of the characteristic, which would be insufficient for the participant to make an informed response. These issues will now be discussed.

5.1.2.1. The Affective Priming Task (APT) measures shared valence rather than attribution of the characteristic to the image

The first problem that was identified with the APT in its conventional form is that it measures to what extent the photographs and descriptions are the same valence, rather than the extent to which the photographed person appears to be described by each characteristic. In the context of the current research using a conventional APT would have meant asking participants to ignore the facial image displayed and simply classify the characteristic as being positive or negative in meaning. The extent to which the face facilitated or inhibited participants response time would, in effect, show the extent to which the face was regarded as being the same valence as the characteristic. This problem could be solved if participants were instructed to answer the question; "do you think the characteristic displayed is an accurate description of the person in the previous photograph?" rather than instructed to ignore the photograph and answer the question "is the characteristic displayed positive or negative in meaning?". It would be preferable to show the photograph before the word because participants are then judging whether a face appeared to have a certain trait. If the word appeared first or simultaneously with the photograph then participants might try to fit the description to the face rather than giving their initial impression.

However, by asking participants to decide whether the characteristic appeared to describe the person, the task would cease to be a truly automatic judgement, and therefore may not be a true APT. In fact the prime would no longer be a prime at all in the sense that it is something that comes before the target; the photograph would be part of the target judgement rather than a prime to the target judgement. Yet this might not be a problem. Klauer et al. (1997), Klauer & Stern (1992), and Wentura (2000) based their explanation of the APT – the affective-matching mechanism (Section 1.4.2.2) - on the finding that in a typical priming task people tend to automatically evaluate whether an adjective (target word) describes a noun (prime word) i.e. they spontaneously ask themselves the question "is object X adjective Y?" when presented with two words. If this is the case then it is certainly feasible that people would automatically ask the question "does this characteristic describe the person in the photograph?" especially given Fazio's comment that images were "likely to produce much more activation from memory" than words (Fazio, 1993a).

More importantly, Bassili (1996) found response latency could serve as an effective implicit measure of attitude strength even if it was not used as part of an APT, IAT or any other task involving a manipulation of response latency. He simply timed participants responses to questions in a telephone interview and found that their response latencies formed part of a construct he termed the operative index of attitude strength. The operative index was based on measures of the psychological process that were not self-reported and was separate from the meta-attitudinal index of attitude strength (which refers to participants' self-knowledge of their own attitudes, and which was formed of self-reported attitudes). Not only did the operative index account for unique variance in attitude pliability (how easy it was to change participants minds with a counter-argument) and attitude stability, but within the operative index response latency was generally more effective at accounting for unique variance than the other measures (the meta-attitudinal index did not account for any unique variance). Therefore it may not be necessary to stick rigidly to the established format of the APT.

Other studies have also attempted to use response latency to measure attitudes to social judgements. For example, Dovido et al. (1986) successfully measured participants' judgements about different racial groups in a priming task, although Judd et al. (1995) failed to replicate their findings. However, assessing the strength of social judgements using response latency does create the problem of multiple responses, which is discussed in the following Section (5.1.2.2).

5.1.2.2. The problem of two response outcomes

The second problem with using the APT was that, in the context of the current research, there was no correct response, i.e. participants could respond "yes" or "no" to the question "does this characteristic describe the previous face?" with equal validity. Fazio (1990b) said this should be avoided. It means that each response now has two outcomes, the direction of the response –"yes" or "no"- and the response latency. This made the study more complicated to analyse.

Dovido et al. (1986) avoided this issue by designing their study in such a way that only one response could be regarded as correct. Their prime words were groups of people, either "white" or "black"; these were followed by adjectives.

They asked participants to either press a “yes” or “no” key to answer the question could the word “ever be true” of the group of people shown previously (i.e. black or white), or was it “always false”. They expected (and largely found) that all answers would be “yes”, i.e. it could be true of a white/black person. To avoid participants pressing the same key in every trial some of the trials were not related to people, in some filler trials the primes were objects, and the adjectives were materials. So participants would be required to respond “no” sometimes, e.g. if the prime is “white” people and the target is “metallic”, then clearly no person can be metallic, just as no tree could be ambitious.

However, this method was not desirable in the present study since we were interested in social judgements related to a relatively small area of a person, their teeth. Including judgements about other objects could distract participants from the details of the photographed people and allow them to respond without ever considering the individuals teeth, because the task would have become matching people with human descriptions and matching objects with the material they are made from. The goal of the study required participants to make careful social judgements not perform a simple matching task. Therefore, a more complicated analysis was considered preferable to a relatively simple analysis of an irrelevant task.

One further advantage of having response direction and response latency as an outcome was that the test effectively became a simultaneous explicit and implicit measure. Response direction was a measure of participants' reported opinions (an explicit test), whereas response latency was a measure of the speed of their decision, i.e. the strength of their opinion towards what they saw. This meant that even if the methodology was not sensitive enough to show different response latencies between the different dental appearances, participants' response direction - whether they said "yes" or "no" to each description for each level of fluorosis - could still be analysed.

5.1.2.3. The problem of the Stimulus Onset-Asynchrony (SOA)

The third and final difficulty with using a traditional APT in the current research was that a typical priming task has Stimulus Onset-Asynchrony (SOA) of 300ms or less (Fazio et al., 1986). In terms of the present study this means there would be just 300ms between the photograph appearing and the description replacing it on the screen. Such a short time is probably insufficient to allow any social impression to form; indeed, an SOA of 300ms is designed to preclude any deliberation or meaningful judgement, which was necessary for the present research goal. Showing photographs of faces probably requires a greater SOA than 300ms, and considering that it was necessary to give participants an opportunity to notice relatively minor variations in the colour of the teeth then a longer SOA seemed absolutely essential.

There are a number of APTs that have successfully used a SOA of longer than 300ms when utilising stimuli that are more complex than typed words. Fazio & Dunton (1997), Fazio et al. (1995), and Towles-Schwen & Fazio (2003) conducted studies with the bona fide pipeline (an APT they adapted to measure racial prejudice) in which the faces were displayed for 315ms followed by a gap of 135ms where the screen was blank, which gave a total SOA of 450ms. However, the bona fide pipeline is used to measure racial attitudes, and a person's race is likely to be more salient than small variations in their tooth colour, so 315ms was still not considered to be a sufficient display time for the photographs in the current study. Hermans et al. (1998) used odours as the prime in an APT. Allowing enough time for the odour to reach the participants nostrils, meant using a SOA of around 10 seconds, which seemed excessive for this research. Bearing in mind the present research involved the participants making social judgements it seemed advisable to follow the lead of the Dovido et al. (1986) study, which also involved social judgements. They displayed the prime for 2000ms (although as was described above it wasn't a prime in the usual sense because it formed part of the judgement), and then they left a gap of 500ms in which the screen was blank, giving a total SOA of 2,500ms. A display time of 2000ms seemed like a reasonable display time for the photographs in the current study, and a gap of 500ms was used to orientate participants to where the word would appear on the screen.

Therefore, given the need to inform participants about fluorosis and the time requirements that would be needed to use the IAT, it was decided that a modified APT, using a SOA of 2500ms, in which participants made social judgements, was the best implicit measure for the current study. Having described the rationale of the methodology, its success as an implicit attitude measure will now be discussed.

5.2. The success of the methodology used in this research

The current methodology had two outcomes, response direction (whether participants pressed the "yes" or "no" key) and response latency (how long the participants took to respond). Response direction can be regarded as an explicit measure of attitude since participants were aware they were responding to a question. Whereas response latency was intended to be an implicit measure of their attitude, since participants were unaware their response latencies were being recorded.

Participants' response direction revealed that social judgements about a person changed when that person had severe fluorosis or dental caries. Participants' response latencies revealed that the strength of participants' attitudes also changed when rating severe fluorosis or dental caries, such that their opinions were weaker when they made a complimentary response (a "yes" response to a positive characteristic or a "no" response to a negative characteristic) and stronger when making an uncomplimentary response (a "no" response to a positive characteristic or a "yes" response to a negative characteristic). The current methodology was therefore successful, both as an explicit measure and an implicit measure.

After a thorough search of the literature (Appendix A) it appears that this is the first time an implicit measure of attitude strength has been used in the assessment of dental appearance. This type of attitude measurement may be useful in future research for assessing attitudes to other dental conditions (or other investigations into altered appearance) and could possibly help to inform decisions about whether or not a particular treatment should be administered, or whether or not a particular appearance is undesirable. Although it could be argued that, in the present research, the strength of participants' attitudes have not revealed differences between the various dental appearances that were not evident from their (more easily measured) direction of response, this may not be the case in other fields of research.

The research discussed in Section 1.4 showed that many previous studies that utilised an implicit measure of attitude were assessing prejudice (particularly racial prejudice; Cunningham et al., 2001; Fazio et al., 1995; Greenwald et al., 1998; Rudman et al., 1999). It would be advantageous to apply the current methodology to a field such as racial prejudice in order to provide validation for it. Furthermore, since the current methodology can be used to determine particular characteristics attributed to the photographed faces it may be able to identify the assumptions that underlie prejudice, rather than simply revealing a global preference of one race over another, which has been the approach taken by the IAT and APT. This is one potential avenue for future research.

5.3. Issues and limitations of the current research

5.3.1. The use of photographs

The findings described in this thesis were generated by asking participants to make judgements based on static, simple photographs. While the use of extra-oral images may be preferable to intra-oral images they may not reflect the reality of decision making in more complex social and personal environments. Future research could therefore explore the possibility of using a prosthesis to create the appearance of various dental conditions on the teeth of an actor, who could then either directly interact with participants in a more life-like social interaction, or could be filmed socially interacting to create video footage that could be rated by participants. This method may prove problematic in terms of successfully recreating a dental condition with a prosthesis, but studies by Berry & Miller (2001) and Asendorpf et al. (2002) assessed the impact of physical attractiveness (though not with emphasis on dental appearance) and shyness using staged social interactions, so there is a methodological precedent to assess social interactions.

5.3.2. Participants multiple ratings of the same target face

The methodology in all three of the experiments in this thesis involved participants rating different dental appearances within the same face (or intra-oral image in the case of the first experiment). In other studies that have manipulated a photograph to display numerous dental appearances within the same face it has been typical to show the participants only one image of the background face (Eli et al., 2001; Feng et al., 2001; Newton et al., 2003; Shaw, 1981; Shaw & Humphreys, 1982; Shaw et al., 1985). This has the advantage that participants are more likely to assume the image is an unaltered photograph. Moreover, the second question sheet administered in the final study revealed that the majority of participants (35 out of 40) assumed that the images of severe fluorosis were a deliberate manipulation of the photograph. However, this does not appear to have had a great impact on their responses, since they clearly indicated that they regarded the images of severe fluorosis less favourably than images of the same face with normal enamel. This would suggest they responded as if they believed each face was a genuine photograph. Furthermore, the current research sought to identify different attitudes to various dental conditions that were only attributable to dental appearance. If the various dental appearances had been judged by different participants then this would have allowed for variation attributable to the participant, rather than just the image. It was for this reason that the same participants were asked to rate different dental appearances within the same face.

5.3.3 The standardisation of images

Given the issue discussed in Section 4.4 regarding the use of digitally manipulated images and the arbitrary nature of creating a representation of a particular dental appearance, perhaps future research should seek to develop a bank of standardised images, approved by a panel of expert judges, for use in future research into dental appearance. Such images could represent a range of dental conditions shown within the same background face or intra-oral image. The background faces could include both males and females at a range of ages, and of various levels of physical attractiveness and from different races. A bank of photographs such as this would mean future research could attempt to determine if dental appearance is of particular importance at a certain age, or for a certain gender, and could untangle the relationship between facial attractiveness and dental appearance. It would also mean that future research into the social importance of dental appearance could be conducted from the same point of reference, with regard to the images used. This may eventually inform decisions about the availability of cosmetic dental treatment provided by the National Health Service or other health care commissioners.

5.3.4. The participants

The volunteers in the second and third studies were students and therefore clearly may not be representative of the general public. However, due to the length of time the study took to complete (approximately one hour per volunteer) it would have been difficult to recruit members of the public. Therefore the studies were conducted using student volunteers.

Another limitation of the participant sample used in the second and third studies of this thesis was the imbalance in the number of males and females; there were more females than males in both of these studies. Although we attempted to recruit an equal number of males and females this was not possible due to the low response rate of men. However, the findings of the second and third studies were supported to some degree by those of the first study, since similar characteristics were identified as being relevant to dental appearance in all three studies. In the first study the participants were drawn from a wider cross section of society, and contained an equal number of males and females. This suggests that the responses of the participants in the second and third studies are not likely to have been very different to responses from the general population.

5.4. Conclusions

The research in this thesis suggests that when participants are explicitly asked to attribute characteristics to photographed individuals, mild and moderate fluorosis are not perceived differently to normal enamel, when viewed in the context of an extra-oral photograph. However, severe fluorosis and dental caries were judged less favourably than normal enamel. Furthermore, participants attributed characteristics to others that went beyond the aesthetic, based on dental appearance. The present research also found that the strength of participants' attitudes to different dental conditions could be implicitly measured, and that when this was done, the strength of participants' attitudes were not different when making social judgements about somebody with normal enamel, mild fluorosis, or moderate fluorosis, but did change unfavourably when rating severe fluorosis or dental caries.

These findings indicate that previous research which has found mild and moderate fluorosis to be an aesthetic problem may have been caused by emphasizing the appearance of the teeth, by asking participants to rate either intra-oral images or the dental appearance of patients. Future research into the social impact of dental appearance may benefit from the use of standardised extra-oral images that show different dental appearances within the same face, especially when comparing two different conditions to identify which is preferable. The development of a widely available bank of such images would be advantageous in allowing different researchers to use the same images. Since the present research found mild fluorosis to be perceived more favourably than untreated dental caries, one obvious course for future research would be to compare fluorosis with treated dental caries. It may also be beneficial to recreate different dental appearances in an actor using a prosthesis, in order to gain further insight into the effect of dental appearance on social interactions.

Another potential direction for future research is to utilise the methodology described in this thesis to implicitly measure attitude strength to other dental conditions, or apply it to other investigations of physical appearance. It could also seek to validate the use of the current methodology by employing it in a research area where implicit measures are established, such a racial prejudice.

Bibliography

Alkhatib, M. N., Holt, R., & Bedi, R. (2004). Aesthetically objectionable fluorosis in the United Kingdom. *British Dental Journal*, **197,** 325-8.

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, **9,** 272-9.

Arndt, J., Greenberg, J., & Cook, A. (2002). Mortality salience and the spreading activation of worldview-relevant constructs: Exploring the cognitive architecture of terror management. *Journal of Experimental Psychology: General*, **131,** 307-24.

Asendorpf, J. B., Banse, R., & Mucke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, **83,** 380-93.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 145-60.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996a). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, **32,** 104-28.

Bargh, J. A., Chen, M., & Burrows, L. (1996b). Automaticity of social behaviour: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, **71,** 230-44.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, **126,** 925-45.

Barsden, A. (1999). "Risk periods" associated with the development of dental fluorosis in maxillary permanent central incisors: a meta-analysis. *Acta Odontologica Scandinavica*, **57,** 247-56.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, **71,** 637-53.

Beltran-Aguilar, E. D., Griffin, S. O., & Lockwood, S. A. (2002). Prevalence and trends in enamel fluorosis in the United States from the 1930s to the 1980s. *Journal of the American Dental Association*, **133,** 157-65.

Berry, D. S., & Miller, K. M. (2001). When boy meets girl: Attractiveness and the five-factor model in opposite-sex interactions. *Journal of Research in Personality*, **35,** 62-77.

Bowen, W. H. (2002). Fluorosis: Is it really a problem? *Journal of the American Dental Association*, **133,** 1405-7.

Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, **81,** 760-73.

Carr, A., Harris, D., & James, C. (2000). The derriford appearance scale (DAS-59): A new scale to measure individual responses to living with problems of appearance. *British Journal of Health Psychology*, **5,** 201-15.

Castelli, L., Zogmaister, C., Smith, E. R., & Arcuri, L. (2004). On the automatic evaluation of social exemplars. *Journal of Personality and Social Psychology*, **86,** 373-87.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behaviour link and social interaction. *Journal of Personality and Social Psychology*, **76,** 893-910.

Chikte, U. M., Louw, A. J., & Stander, I. (2001). Perceptions of fluorosis in Northern Cape communities. *South African Dental Journal*, **56,** 528-32.

Clark, D. C. (1994). Trends in prevalence of dental fluorosis in North America. *Community Dentistry and Oral Epidemiology*, **22,** 148-52.

Clark, D. C. (1995). Evaluation of aesthetics for the different classifications of the tooth surface index of fluorosis. *Community Dentistry and Oral Epidemiology*, **23,** 80-3.

Clark, D. C., Hann, H. J., Williamson, M. F., & Berkowitz, J. (1993). Aesthetic concerns of children and parents in relation to different classifications of the tooth surface index of fluorosis. *Community Dentistry and Oral Epidemiology*, **21,** 360-4.

Clifford, M. M., & Walster, E. (1973). Research note: The effect of physical attractiveness on teacher expectations. *Social Education*, **46,** 248-58.

Cons, N. C., Jenny, J., Kohout, F., J., Freer, T. J., & Eismann, D. (1983). Perceptions of occlusal conditions in Australia, the German Democratic Republic and the United States of America. *International Dental Journal*, **33,** 200-6.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology*, **85,** 639-49.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, **12,** 163-70.

Dare, G. J. (1992). The effect of pupil appearance on teacher expectations. *Early Child Development and Care*, **80,** 97-101.

Dasgupta, N., McGhee, D. B., Greenwald, A. G., & Banaji, M. R. (2001). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology*, **36,** 316-28.

De Houwer, J. (2001). A structural and process analysis of the implicit association test. *Journal of Experimental Social Psychology*, **37,** 443-51.

De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. *In* J. Musch, & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219-44). Mahwah: Lawrence Erlbaum Associates.

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective and identity priming with episodically associated stimuli. *Cognition and Emotion*, **12,** 145-69.

de Jong, P. J. (2002). Implicit self-esteem and social anxiety: Differential self-favouring effects in high and low anxious individuals. *Behaviour Research and Therapy*, **40,** 501-8.

de Jong, P. J., Pasman, W., Kindt, M., & van den Hout, M. A. (2001). A reaction time paradigm to assess (implicit) complaint-specific dysfunctional beliefs. *Behaviour Research and Therapy*, **39,** 101-13.

Dean, H. T. (1934). Classification of mottled enamel diagnosis. *Journal of the American Dental Association*, **21**, 1421-6.
Demelleweek, C., Humphris, G. M., Hare, M., & Brown, J. (1997). Children's perception of, and attitude towards, unfamiliar peers with facial port-wine stains. *Journal of Pediatric Psychology*, **22**, 471-85.
Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behaviour or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology*, **74**, 865-77.
Dion, K. K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, **24**, 285-90.
Dovido, J. F., Evans, N., & Tyler, R. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, **22**, 22-37.
Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, **13**, 513-19.
E.T.S.L. (2001). Benchmarking results. Experimental Timing Standards Laboratory. http://www.psychology.ltsn.ac.uk/ETSL/. 10/8/2004.
Eagly, A. H., Makhijani, M. G., Ashmore, R. D., & Longo, L. C. (1991). What is beautiful is good, but...: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, **110**, 109-28.
Eli, I., Bar-Tal, Y., & Kostovetzki, I. (2001). At first glance: Social meanings of dental appearance. *Journal of Public Health Dentistry*, **61**, 150-4.
Ellwood, R. P., & O'Mullane, D. (1995). Enamel opacities and dental esthetics. *Journal of Public Health Dentistry*, **55**, 171-6.
Evans, R. W., & Darvell, B. W. (1995). Refining the estimate of the critical period for susceptibility to enamel fluorosis in human maxillary central incisors. *Journal of Public Health Dentistry*, **55**, 238-49.
Evans, R. W., & Stamm, J. W. (1991). An epidemiologic estimate of the critical period during which human maxillary central incisors are most susceptible to fluorosis. *Journal of Public Health Dentistry*, **51**, 251-9.
Everett, E. T., McHenry, M. A. K., Reynolds, N., Eggertsson, H., Sullivan, J., Kantmann, C., Martinez- Mier, E. A., Warrick, J. M., & Stookey, G. K. (2002). Dental fluorosis: Variability among different inbred mouse strains. *Journal of Dental Research*, **81**, 794-8.
Fazio, R. H. (1990a). Multiple processes by which attitudes guide behaviour: The MODE model as an integrative framework. *In* M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 75-109). San Diego, CA.: Academic Press.
Fazio, R. H. (1990b). A practical guide to the use of response latency in social psychological research. *In* C. Hendrick, & M. S. Clark (Eds.), *Review of personality and social psychology* (pp. 74-97). Newbury Park, C.A.: Sage.
Fazio, R. H. (1993a). Addendum: Reply to Chaiken and Bargh by Russell H. Fazio. *Journal of Personality and Social Psychology*, **64**, 764-5.
Fazio, R. H. (1993b). Variability in the likelihood of automatic attitude activation: Data reanalysis and commentary on Bargh, Chaiken, Govender, and Pratto (1992). *Journal of Personality and Social Psychology*, **64**, 753-8.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. *In* R. E. Petty, & J. A. Krosnick (Eds.), *Attitude strength - antecedents and consequences* (pp. 247-83). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, **15,** 115-41.

Fazio, R. H., Chen, J., McDonel, E. C., & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology*, **18,** 339-57.

Fazio, R. H., & Dunton, B. C. (1997). Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology*, **33,** 451-70.

Fazio, R. H., & Hilden, L. E. (2001). Emotional reactions to a seemingly prejudiced response: The role of automatically activated racial attitudes and motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, **27,** 538-49.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, **69,** 1013-27.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, **54,** 297-328.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, **50,** 229-38.

Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behaviour processes. *In* S. Chaiken, & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 97-116). New York: Guilford Press.

Fazio, R. H., Williams, C. J., & Powell, M. C. (2000). Measuring associative strength: Category item associations and their activation from memory. *Political Psychology*, **21,** 7-25.

Feingold, A. (1992a). Gender differences in mate selection preferences: A test of the parental investment model. *Psychological Bulletin*, **112,** 125-39.

Feingold, A. (1992b). Good-looking people are not what we think. *Psychological Bulletin*, **111,** 304-41.

Feng, X. P., Newton, J. T., & Robinson, P. G. (2001). The impact of dental appearance on perceptions of personal characteristics among Chinese people in the United Kingdom. *International Dental Journal*, **51,** 282-6.

Ferguson, M. J., & Bargh, J. A. (2003). The constructive nature of automatic evaluation. *In* J. Musch, & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 169-89). Mahwah: Lawrence Erlbaum Associates.

Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences*, **8,** 33-9.

Ferguson, M. J., Bargh, J. A., & Nayak, D. (in press). After-affects: How automatic attitudes influence the interpretation of subsequent, unrelated stimuli. *Journal of Experimental Psychology: General.*

Fiske, J., Gelbier, S., & Watson, R. M. (1990). The benefit of dental care to an elderly population assessed using a sociodental measure of oral handicap. *British Dental Journal*, **168,** 153-6.

Garcia, M. T., & Bargh, J. A. (2003). Automatic evaluation of novel words: The role of superficial phonetics. *Journal of Language and Social Psychology*, **22,** 414-33.

Giner-Sorolla, R., Garcia, M. T., & Bargh, J. A. (1999). The automatic evaluation of pictures. *Social Cognition*, **17,** 76-96.

Grammer, K., & Thornhill, R. (1994). Human (Homo sapiens) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, **108,** 233-42.

Gray, N. S., MacCulloch, M. J., Smith, J., Morris, M., & Snowden, R. J. (2003). Violence viewed by psychopathic murderers: Adapting a revealing test may expose those psychopaths who are most likely to kill. *Nature*, **423,** 497-8.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, **273,** 1699-702.

Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self concept. *Journal of Personality and Social Psychology*, **79,** 1022-38.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, **74,** 1464-80.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the implicit association test at age 3. *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 85-93.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, **85,** 197-216.

Griffin, S. O., Beltran, E. D., Lockwood, S. A., & Barker, L. K. (2002). Esthetically objectionable fluorosis attributable to water fluoridation. *Community Dentistry and Oral Epidemiology*, **30,** 199-209.

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, **9,** 139-50.

Hall, D., & Hall, I. (1996). *Practical social research*. Basingstoke: Macmillan.

Hawley, G. M., Ellwood, R. P., & Davies, R. M. (1996). Dental caries, fluorosis and the cosmetic implications of different TF scores in 14 year old adolescents. *Community Dental Health*, **13,** 189-92.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, **82,** 903-18.

Hermans, D., Baeyens, F., & Eelen, P. (1998). Odours as affective-processing context for word evaluation: A case of cross-modal affective priming. *Cognition and Emotion*, **12,** 601-13.

Hermans, D., Crombez, G., & Eelen, P. (2000). Automatic attitude activation and efficiency: The fourth horseman of automaticity. *Psychologica Belgica*, **40,** 3-22.

Hermans, D., De Houwer, J., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition and Emotion*, **8,** 515-33.

Himmelfarb, S. (1993). The measurement of attitudes. *In* A. H. Eagly, & S. Chaiken (Eds.), *The psychology of attitudes*. Orlando: Harcourt Brace Jovanovich College Publishers.

Holloway, P. J., & Ellwood, R. P. (1997). The prevalence, causes and cosmetic importance of dental fluorosis in the United Kingdom: A review. *Community Dental Health*, **14,** 148-55.

Horowitz, H. (1992). The need for toothpastes with lower than conventional fluoride concentrations for preschool-aged children. *Journal of Public Health Dentistry*, **52,** 216-21.

Horowitz, H. S., Driscoll, W. S., Meyers, R. J., Heifetz, S. B., & Kingman, A. (1984). A new method for assessing the prevalence of dental fluorosis - the tooth surface index of fluorosis. *Journal of the American Dental Association*, **109,** 37-41.

Jackson, D. J., & Huston, T. L. (1975). Physical attractiveness and assertiveness. *Journal of Social Psychology*, **96 (First Half),** 79-84.

Judd, C. M., Park, B., Ryan, C. S., Brauer, M., & Kraus, S. (1995). Stereotypes and ethnocentrism: Diverging interethnic perceptions of African American and White American youth. *Journal of Personality and Social Psychology*, **69,** 460-81.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, **81,** 774-88.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. *In* J. Musch, & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7-49). Mahwah: Lawrence Erlbaum Associates.

Klauer, K. C., RoBnagel, C., & Musch, J. (1997). List context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23,** 246-55.

Klauer, K. C., & Stern, E. (1992). How attitudes guide memory-based judgments: A two process model. *Journal of Experimental Social Psychology*, **28,** 186-206.

Kuehnen, U., Schiessl, M., Bauer, N., Paulig, N., Poehlmann, C., & Schmidthals, K. (2001). How robust is the IAT? Measuring and manipulating implicit attitudes of East- and West-Germans. *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 135-44.

Lalumandier, J. A., & Rozier, G. (1998). Parents' satisfaction with children's tooth color: Fluorosis as a contributing factor. *Journal of the American Dental Association*, **129,** 1000-6.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, **126,** 390-423.

Lansdown, R., Lloyd, J., & Hunter, J. (1991). Facial deformity in childhood: Severity and psychological adjustment. *Child: Care, Health and Development*, **17,** 165-71.

Lesan, W. R. (1987). Dental fluorosis: A review of literature with comments on tropical characteristics. *East African Medical Journal*, **64,** 493-8.

Levy, S. M., & Guha-Chowdhury, N. (1999). Total fluoride intake and implications for dietry fluoride supplementation. *Journal of Public Health Dentistry*, **59,** 211-23.

Levy, S. M., Hillis, S. L., Warren, J. J., Broffitt, B. A., Mahbubul Islam, A. K. M., Wefel, J. S., & Kanellis, M. J. (2002). Primary tooth fluorosis and fluoride intake during the first year of life. *Community Dentistry and Oral Epidemiology*, **30,** 286-95.

Levy, S. M., Kiritsy, M. C., & Warren, J. J. (1995). Sources of fluoride intake in children. *Journal of Public Health Dentistry*, **55,** 39-52.

Levy, S. M., Warren, J. J., & Jakobsen, J. R. (2002). Follow-up study of dental students' esthetic perceptions of mild dental fluorosis. *Community Dentistry and Oral Epidemiology*, **30,** 24-8.

Lister, I. (2001). *The psychology of facial disfigurement: A guide for health and social care professionals*. London: Changing Faces.

Livingston, R. W. (2001). What you see is what you get: Systematic variability in perceptual-based social judgment. *Personality and Social Psychology Bulletin*, **27,** 1086-96.

Macgregor, F. C. (1970). Social and psychological implications of dentofacial disfigurement. *Angle Orthodontist*, **40,** 231-3.

Maner, J. K., Kenrick, D. T., Becker, D. V., Delton, A. W., Hofer, B., Wilbur, C. J., & Neuberg, S. L. (2003). Sexually selective cognition: Beauty captures the mind of the beholder. *Journal of Personality and Social Psychology*, **85,** 1107-20.

Marsh, H. W., Hey, J., Roche, L. A., & Perry, C. (1997). Structure of physical self-concept: Elite athletes and physical education students. *Journal of Educational Psychology*, **89,** 369-80.

Marsh, K. L., Johnson, B. T., & Scott-Sheldon, L. A. (2001). Heart versus reason in condom use: Implicit versus explicit attitudinal predictors of sexual behavior. *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 161-75.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. *In* J. F. Dovidio, & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91-125). Orlando: Harcourt Brace Jovanovich.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in America? It depends on who is asking and what is asked. *Journal of Conflict Resolution*, **25,** 563-79.

McCrae, R. R., & Costa Jr., P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, **52,** 509-16.

McDonagh, M., Whiting, P., Bradley, M., Cooper, J., Sutton, A., Chestnutt, I., Misso, K., Wilson, P., Treasure, E., & Kleijnen, J. (2000). A systematic review of water fluoridation. Report 18 NHS centre for reviews and dissemination (pp. 34-45). York: The University of York.

McKnight, C. B., Levy, S. M., Cooper, S. E., & Jakobsen, J. R. (1998). A pilot study of esthetic perceptions of dental fluorosis vs. selected other dental conditions. *Journal of Dentistry for Children*, **65,** 233-8.

McKnight, C. B., Levy, S. M., Cooper, S. E., Jakobsen, J. R., & Warren, J. J. (1999). A pilot study of dental students' esthetic perceptions of computer-generated mild dental fluorosis compared to other conditions. *Journal of Public Health Dentistry*, **59,** 18-23.

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 107-22.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the implicit association test. *Journal of Personality and Social Psychology*, **85,** 1180-92.

Miller, A. G. (1970). Role of physical attractiveness in impression formation. *Psychonomic Science*, **19,** 241-3.

Milsom, K., & Mitropoulos, C. M. (1990). Enamel defects in 8-year-old children in fluoridated and non-fluoridated parts of Cheshire. *Caries Research*, **24,** 286-9.

Milsom, K. M., Tickle, M., Jenner, A., & Peers, A. (2000). A comparison of normative and subjective assessment of the child prevalence of developmental defects of enamel amongst 12-year-olds living in the north west region. *Public Health*, **114,** 340-4.

Moser, C. A., & Kalton, G. (1972). *Survey methods in social investigation*. London: Heinemann Educational Books Ltd.

Moss, T. (1997). Individual variation in adjusting to visible differences. *In* R. Lansdown, N. Rumsey, E. Bradbury, A. Carr, & J. Partridge (Eds.), *Visibly different: Coping with disfigurement*. Oxford: Butterworth-Heinemann.

Murray, J. J. (2003). Fluorides and dental caries. *In* J. J. Murray, J. H. Nunn, & J. G. Steele (Eds.), *Prevention of oral disease* (pp. 35-60). New York: Oxford University Press.

Murray, J. J., Rugg-Gunn, A. J., & Jenkins, G. N. (1976). *Fluorides in caries prevention*. Oxford: Butterworth-Heinemann.

Musch, J., & Klauer, K. C. (2003). The psychology of evaluation: An introduction. *In* J. Musch, & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 1-5). Mahwah: Lawrence Erlbaum Associates.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, **106,** 225-54.

Newbrun, E., & Horowitz, H. (1999). Why we have not changed our minds about the safety and efficacy of water fluoridation. *Perspectives in Biology and Medicine*, **42,** 526-43.

Newton, J. T., Prabhu, N., & Robinson, P. G. (2003). The impact of dental appearance on the appraisal of personal characteristics. *The International Journal of Prosthodontics*, **16,** 429-34.

Niedenthal, P. M. (1990). Implicit perception of affective information. *Journal of Experimental Social Psychology*, **26,** 505-27.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math not = me. *Journal of Personality and Social Psychology*, **83,** 44-59.

Office for National Statistics (2003). Childrens dental health in the United Kingdom - Non-carious dental conditions: Office for National Statistics.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science*, **14,** 636-9.

Olson, M. A., & Fazio, R. H. (2004a). Reducing the influence of extrapersonal associations on the implicit association test: Personalizing the IAT. *Journal of Personality and Social Psychology*, **86,** 653-67.
Olson, M. A., & Fazio, R. H. (2004b). Trait inferences as a function of automatically activated racial attitudes and motivation to control prejudiced reactions. *Basic and Applied Social Psychology*, **26,** 1-11.
Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement*. London: Pinter Publishers Ltd.
Pallant, J. (2001). *SPSS survival manual*. Maidenhead: Open University Press.
Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, **12,** 729-38.
Plessner, H., & Banse, R. (2001). Attitude measurement using the implicit association test (IAT): Editorial. *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 82-4.
Pratto, F., & John, O. P. (1991). Automatic vigilence: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, **61,** 380-91.
Reber, A. S. (1995). *Dictionary of psychology*. Harmondsworth: Penguin.
Rhodes, G., Zebrowitz, L. A., Clark, A., Kalick, S. M., Hightower, A., & McKay, R. (2001). Do facial averageness and symmetry signal health? *Evolution and Human Behavior*, **22,** 31-46.
Richmond, S., Tatarunaite, E., Playle, R., Hood, K., & Shaw, W. (in press). Facial attractiveness: A longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics*.
Riordan, P. J. (1993a). Perceptions of dental fluorosis. *Journal of Dental Research*, **72,** 1268-74.
Riordan, P. J. (1993b). Specialist clinicians' perceptions of dental fluorosis. *Journal of Dentistry for Children*, **60,** 315-20.
Riordan, P. J. (2002). Dental fluorosis decline after changes to supplement and toothpaste regimens. *Community Dentistry and Oral Epidemiology*, **30,** 233-40.
Ripa, L. W. (1991). A critique of topical fluoride methods (dentifrices, mouthrinses, operator-, and self-applied gels) in an era of decreased caries and increased fluorosis prevalence. *Journal of Public Health Dentistry*, **51,** 23-40.
Roberts-Harry, D. P., Hathorn, I. S., & Stephens, C. D. (1992). The ranking of facial attractiveness. *European Journal of Orthodontics*, **14,** 483-8.
Rodd, H. D., & Davidson, L. E. (1997). The aesthetic management of severe dental fluorosis in the young patient. *Dental Update*, **24,** 408-11.
Roefs, A., & Jansen, A. (2002). Implicit and explicit attitudes toward high-fat foods in obesity. *Journal of Abnormal Psychology*, **111,** 517-21.
Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the implicit association test (IAT). *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 94-106.
Rozier, G. (1999). The prevalence and severity of enamel fluorosis in North American children. *Journal of Public Health Dentistry*, **59,** 239-46.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the implicit association test. *Social Cognition*, **17,** 437-65.

Rumsey, N. (1998). Visible disfigurement. *In* D. W. Johnston, & M. Johnston (Eds.), *Comprehensive clinical psychology: Vol. 8 Health psychology* (pp. 575-93). Oxford: Pergamon.

Rutt, C. D., & Coleman, K. J. (2001). The evaluation of a measurement model for the body image questionnaire and the eating attitudes test in a Hispanic population. *Hispanic Journal of Behavioral Sciences*, **23,** 153-70.

Salvia, J., Algozzine, R., & Sheare, J. B. (1977). Attractiveness and school achievement. *Journal of School Psychology*, **15,** 60-7.

Schuette, R. A., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin*, **21,** 704-10.

Shaw, W. C. (1981). The influence of children's dentofacial appearance on their social attractiveness as judged by peers and lay adults. *Journal of American Orthodontics*, **79,** 399-415.

Shaw, W. C., & Humphreys, S. (1982). Influence of children's dentofacial appearance on teacher expectations. *Community Dentistry and Oral Epidemiology*, **10,** 313-9.

Shaw, W. C., Rees, G., Dawe, M., & Charles, C. R. (1985). The influence of dentofacial appearance on the social attractiveness of young adults. *American Journal of Orthodontics*, **87,** 21-6.

Sigurjons, H., Cochran, J. A., Ketley, C. E., Holbrook, W. P., Lennon, M. A., & O'Mullane, D. M. (2004). Parental perception of fluorosis among 8-year-old children living in three communities in Iceland, Ireland and England. *Community Dentistry and Oral Epidemiology*, **32 Supp/1,** 34-8.

Simpson, J. A., Gangestad, S. W., & Lerma, M. (1990). Perception of physical attractiveness: Mechanisms involved in the maintenance of romantic relationships. *Journal of Personality and Social Psychology*, **59,** 1192-201.

Smith, E. R., Fazio, R. H., & Cejka, M. A. (1996). Accessible attitudes influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology*, **71,** 888-98.

Soto-Rojas, A. E., Martinez- Mier, E. A., Maupome, G., & Urena-Cirett, J. L. (2002). Impact of dental fluorosis on well-being in Mexico City children. *Journal of Dental Research*, **81,** 218.

Steffens, M. C., & Plewe, I. (2001). Items cross-category associations as a confounding factor in the implicit association test. *Zeitschrift Fuer Experimentelle Psychologie*, **48,** 123-34.

Streiner, D. L., & Norman, G. R. (1995). *Health measurement scales: A practical guide to their development and use*. New York: Oxford University Press.

Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (2001). Using the implicit association test to investigate attitude-behaviour consistency for stigmatised behaviour. *Cognition and Emotion*, **15,** 207-30.

Szpunnar, S. M., & Burt, B. A. (1987). Trends in the prevalence of dental fluorosis in the United States: A review. *Journal of Public Health Dentistry*, **47**, 71-9.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.

Tanzer, N. K. (1996). Cross-cultural bias in Likert-type inventories: Perfect matching factor structures and still biased? *European Journal of Psychological Assessment*, **11**, 194-201.

Tavener, J. A., Davies, G. M., Davies, R. M., & Ellwood, R. P. (2004). The prevalence and severity of fluorosis and other developmental defects of enamel in children who received free fluoride toothpaste containing either 440 or 1450 ppm F from the age of 12 months. *Community Dental Health*, **21**, 217-23.

Terry, R. L. (1977). Further evidence on components of facial attractiveness. *Perceptual and Motor Skills*, **45**, 130.

Terry, R. L., & Brady, C. S. (1976). Effects of framed spectacles and contact lenses on self-ratings of facial attractiveness. *Perceptual and Motor Skills*, **42**, 789-90.

Terry, R. L., & Davis, J. S. (1976). Components of facial attractiveness. *Perceptual and Motor Skills*, **42**, 918.

Thompson, A., & Kent, G. (2001). Adjusting to disfigurement: Processes involved in dealing with being visibly different. *Clinical Psychology Review*, **21**, 663-82.

Thurstone, L. L. (1927a). A law of comparative judgement. *Psychological Review*, **34**, 273-86.

Thurstone, L. L. (1927b). Psychophysical analysis. *American Journal of Psychology*, **38**, 368-89.

Thurstone, L. L. (1928). Attitudes can be measured. *Journal of Sociology*, **33**, 529-54.

Towles-Schwen, T., & Fazio, R. H. (2003). Choosing social situations: The relation between automatically activated racial attitudes and anticipated comfort interacting with African Americans. *Personality and Social Psychology Bulletin*, **29**, 170-82.

Trivers, R. (1972). Parental investment and sexual selection. *In* B. Campbell (Ed.), *Sexual selection and the descent of man* (pp. 136-79). London: Heinemann Educational.

Trivers, R. (1985). *Social evolution*. Menlo Park, CA: Benjamin/Cummings.

Urena-Cirett, J. L., Martinez- Mier, E. A., Maupome, G., & Soto-Rojas, A. E. (2002). Impact of dental fluorosis on well-being of children's by their parents perception in a pediatric dental practice in Mexico City. *Journal of Dental Research*, **81**, 218.

van Palenstein Helderman, W. H., & Mkasabuni, E. (1993). Impact of dental fluorosis on the perception of well-being in an endemic fluorosis area in Tanzania. *Community Dentistry and Oral Epidemiology*, **21**, 243-4.

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology*, **82**, 919-34.

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets

reveal evaluative judgment tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26,** 456-69.
Whelton, H. P., Ketley, C. E., McSweeney, F., & O'Mullane, D. M. (2004). A review of fluorosis in the European Union: prevalence, risk factors and aesthetic issues. *Community Dentistry and Oral Epidemiology*, **32 Supp/1,** 9-18.
Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, **72,** 262-74.
Woodward, G. L., Main, P. A., & Leake, J. L. (1996). Clinical determinants of a parent's satisfaction with the appearance of a child's teeth. *Community Dentistry and Oral Epidemiology*, **24,** 416-8.
Yaeger, J. K. (1976). Enamel. *In* S. N. Bhaskar (Ed.), *Orban's oral histology and embryology* (pp. 45-104). Saint Louis: The C.V. Mosby Company.
Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, Monograph Supplement*, **9,** 1-27.
Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998). The relationship between appearance and personality across the life span. *Personality and Social Psychology Bulletin*, **24,** 736-49.
Zebrowitz, L. A., Fellous, J. M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionoist modeling. *Personality and Social Psychology Review*, **7,** 194-215.
Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, **28,** 238-49.
Zebrowitz, L. A., Olson, K., & Hoffman, K. (1993). Stability of babyfaceness and attractiveness across the lifespan. *Journal of Personality and Social Psychology*, **64,** 453-66.
Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the lifespan. *Personality and Social Psychology Bulletin*, **22,** 1258-69.

Appendix A

Literature search

Literature Search

The search for relevant literature involved three strategies, these were search engines, following paper trails, and Zetoc Alerts. These are described below.

Search Engines
Four search engines (MEDLINE, Web of Science, Clin-PSYCH, and PSYC-Info) were used to identify pertinent articles. The literature on the three sub-sections of the review was searched for separately, and each search was conducted several times throughout the course of the PhD (the Fluorosis and Attitude Measure searches were run four times each and the Appearance and Disfigurement search three times). Details of the search terms, search engines, and the dates of the searches are shown below. The findings of the individual search terms were systematically combined in order to reduce the number of articles returned to a manageable size.

Hand Searching
In addition to the search engines, articles and books cited in other work that appeared to be of relevance were located and read.

Zetoc Alerts
The Zetoc Alert facility e-mails the contents page of new editions of selected journals and of new articles containing selected search terms and authors directly to the user. This was utilised throughout the course of this PhD and was especially useful for identifying papers published between or after the electronic searches.

Searches conducted with the search engines

1) Appearance and disfigurement

Search engines
MEDLINE, Web of Science, Clin-PSYCH, PSYC-Info.

Dates of searches
January 2002; July 2003; October 2004.

Search terms
Face; Appearance; Facial appearance; Dental appearance; Appearance preferences; Physical appearance; Attractiveness; Facial attractiveness; Aesthetics; Facial aesthetics; Disfigurement; Body image; Social perception; Self concept; Dental treatment; Gerry Kent; Nichola Rumsey; Jonathon Timothy Newton.

"Appearance" combined with:
Perception; Impression; Attributes; Attributions;
Characteristics; Fluorosis; Disfigurement; Body image; Social perception; Self concept; Dental treatment.

"Attractiveness" combined with:
Perception; Impression; Attributes; Attributions; Characteristics; Fluorosis; Disfigurement; Body image; Social perception; Self concept; Dental treatment.

"Aesthetics" combined with:
Perception; Impression; Attributes; Attributions; Characteristics; Fluorosis; Disfigurement; Body image; Social perception; Self concept; Dental treatment.

"Disfigurement" combined with:
Perception; Impression; Attributes; Attributions; Characteristics; Fluorosis; Body image; Social perception; Self concept; Dental treatment.

2) Fluorosis

Search engines
MEDLINE; Web of Science.

Dates of searches
November 2001; January 2003; July 2003; October 2004.

Search terms
Fluorosis; Face; Appearance; Facial appearance; Dental appearance; Appearance preferences; Physical appearance; Attractiveness; Facial attractiveness; Aesthetics; Facial aesthetics; Dental treatment.

"Fluorosis" combined with:
Physical attractiveness; Attractiveness; Appearance; Aesthetics; Perception; Impression; Attributes; Attributions; Characteristics; Body image; Social perception; Self concept.

3) Attitude measures

Search engines
MEDLINE, Web of Science, Clin-PSYCH, PSYC-Info.

Dates of searches
February 2002; January 2003; November 2003; October 2004.

Search terms
Questionnaire; Questionnaire design; Attitude measure; Implicit attitude; Explicit attitude; Affective Priming; Implicit Association Test; Russell Fazio; John Bargh; Anthony Greenwald; Theory of Reasoned Action; Theory of Planned Behaviour.

Each of these terms was systematically combined with the following:

Appearance; Facial appearance; Dental appearance; Appearance preferences; Physical appearance; Attractiveness; Facial attractiveness;

Aesthetics; Facial aesthetics; Disfigurement; Body image; Social perception; Self concept.

Appendix B

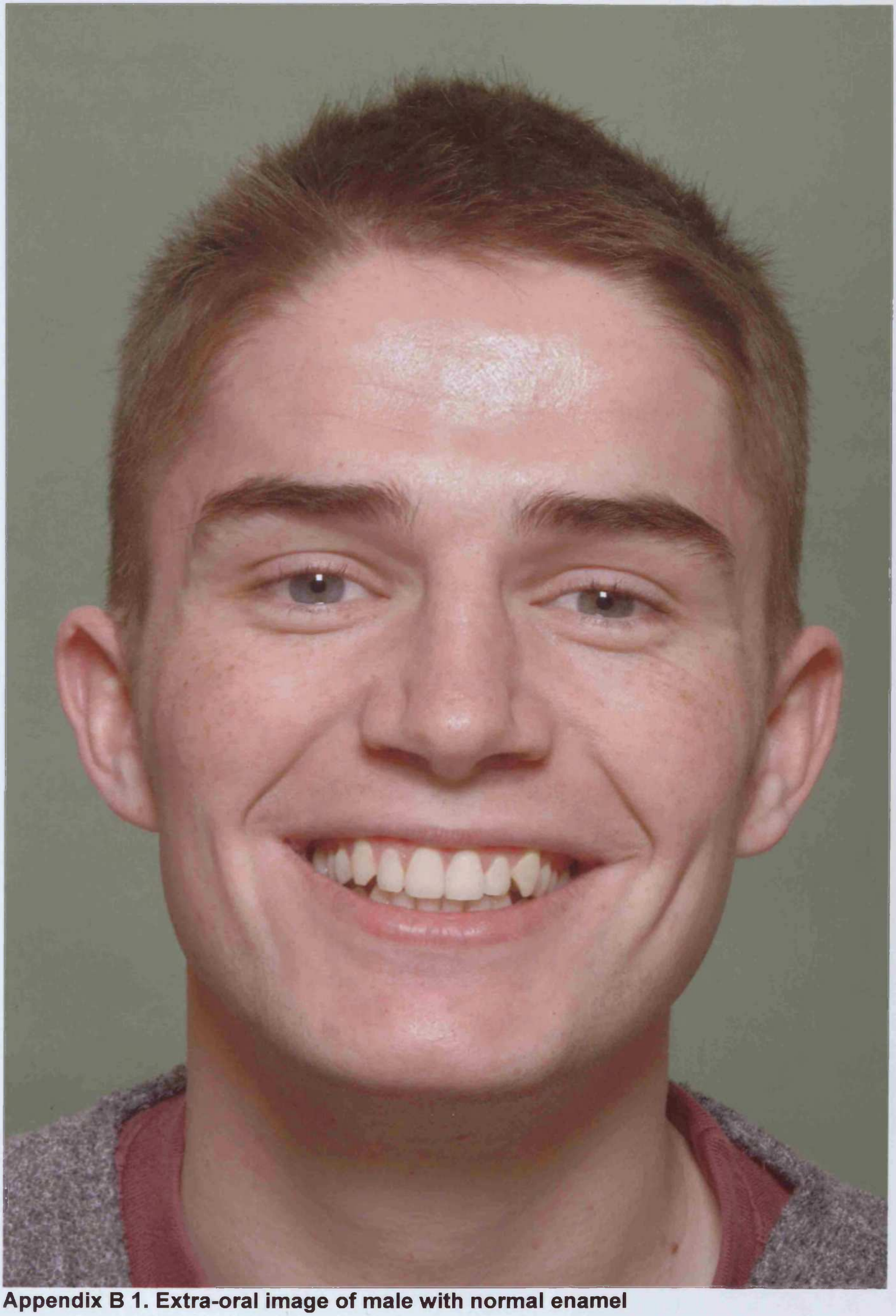
Photographs

- B 1. Extra-oral image of male with normal enamel.
- B 2. Extra-oral image of male with mild fluorosis.
- B 3. Extra-oral image of male with moderate fluorosis.
- B 4. Extra-oral image of male with severe fluorosis.
- B 5. Extra-oral image of male with dental caries.

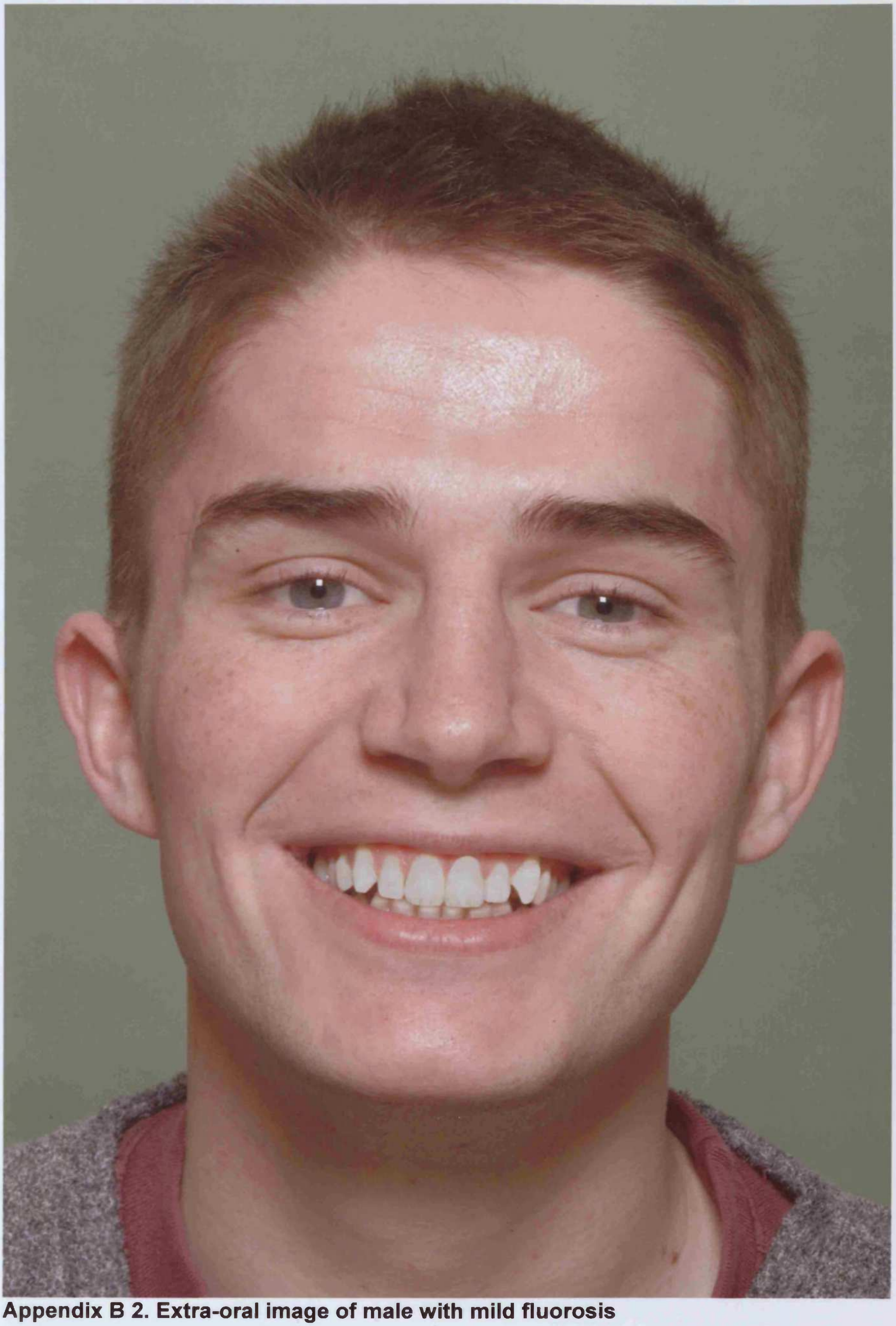
- B 6. Extra-oral image of female with normal enamel.
- B 7. Extra-oral image of female with mild fluorosis.
- B 8. Extra-oral image of female with moderate fluorosis.
- B 9. Extra-oral image of female with severe fluorosis.
- B 10. Extra-oral image of female with dental caries.

- B 11. Intra-oral image of male with normal enamel.
- B 12. Intra-oral image of male with mild fluorosis.
- B 13. Intra-oral image of male with moderate fluorosis.
- B 14. Intra-oral image of male with severe fluorosis.

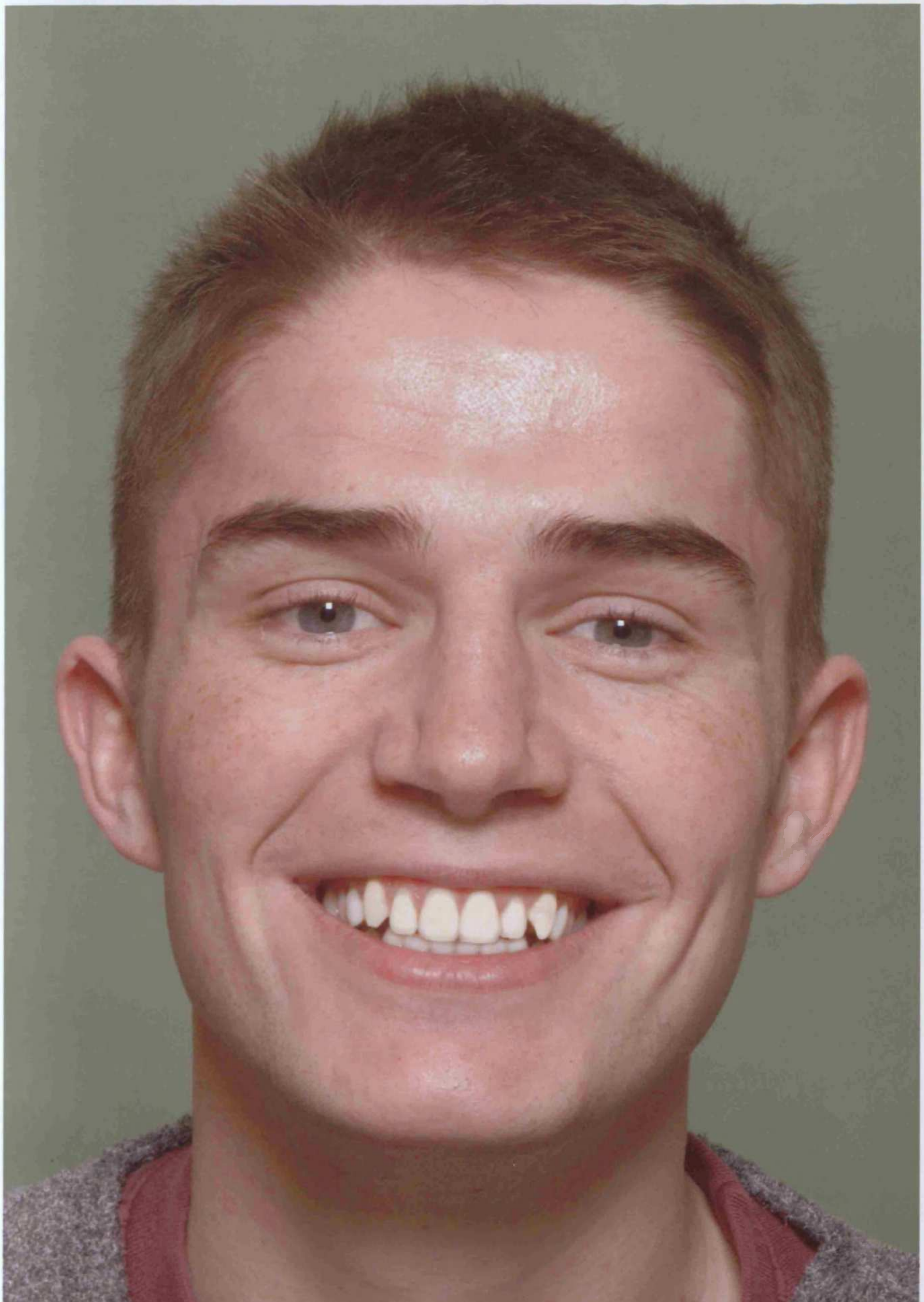
- B 15. Intra-oral image of female with normal enamel.
- B 16. Intra-oral image of female with mild fluorosis.
- B 17. Intra-oral image of female with moderate fluorosis.
- B 18. Intra-oral image of female with severe fluorosis.



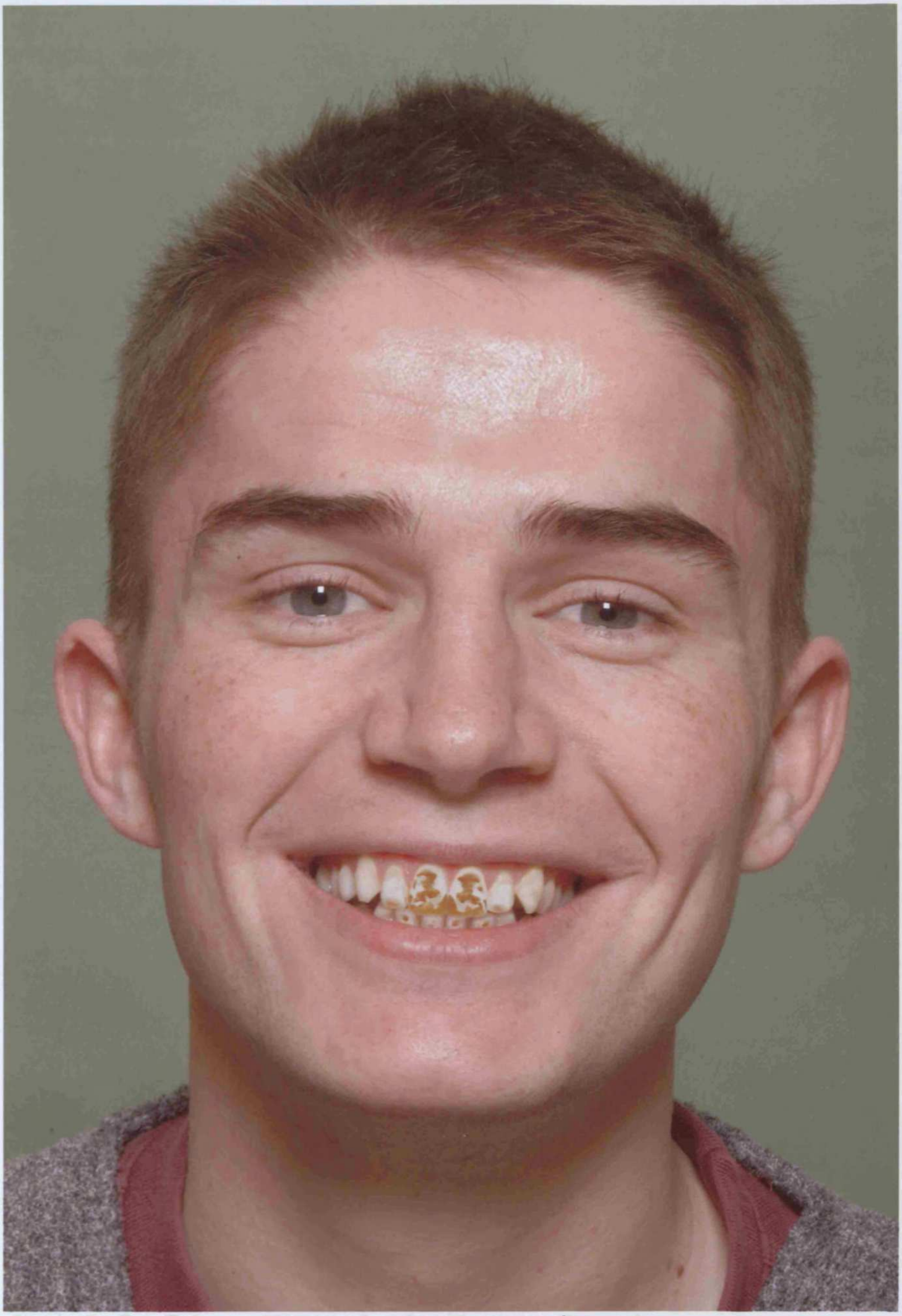
Appendix B 1. Extra-oral image of male with normal enamel



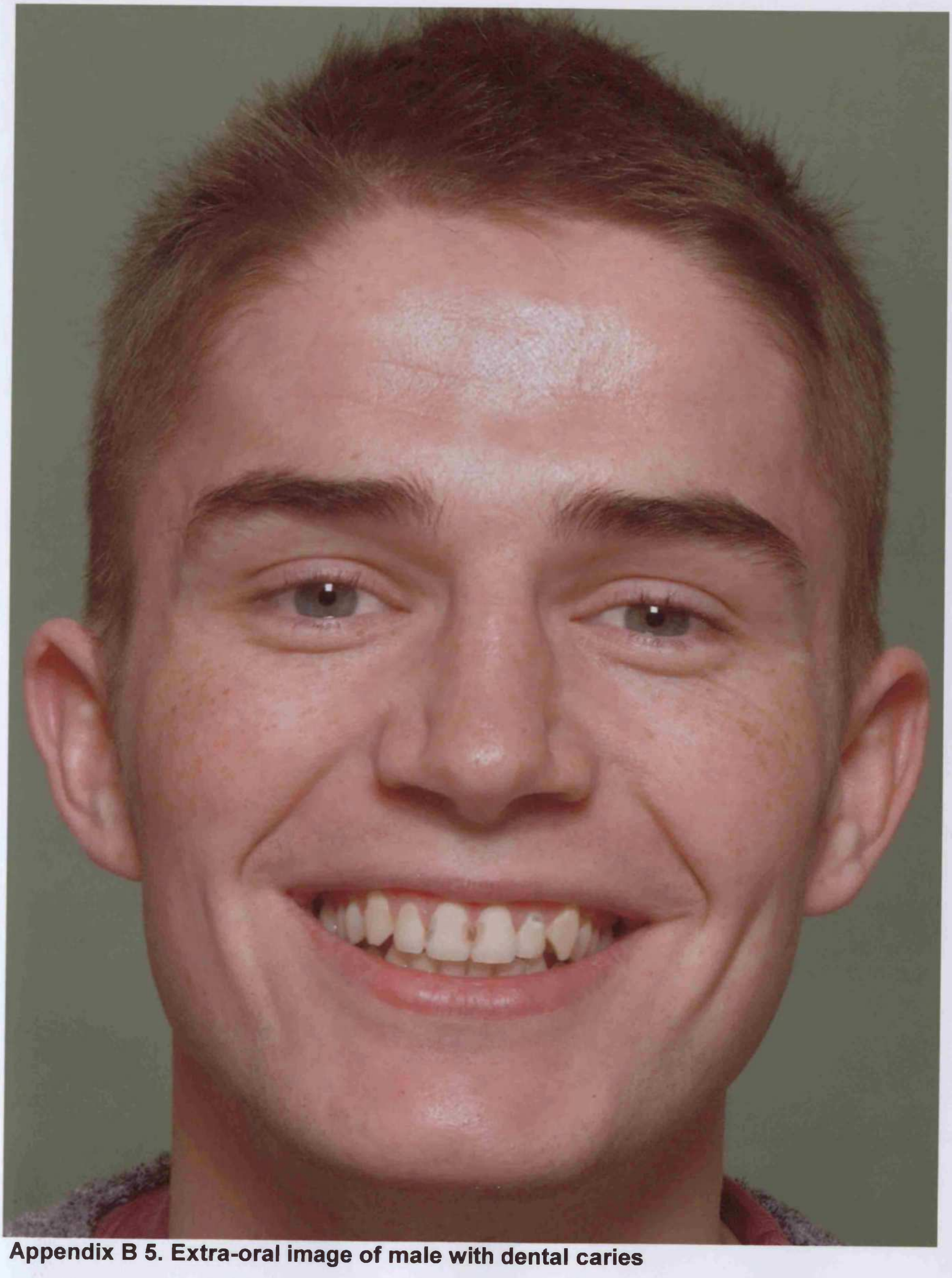
Appendix B 2. Extra-oral image of male with mild fluorosis



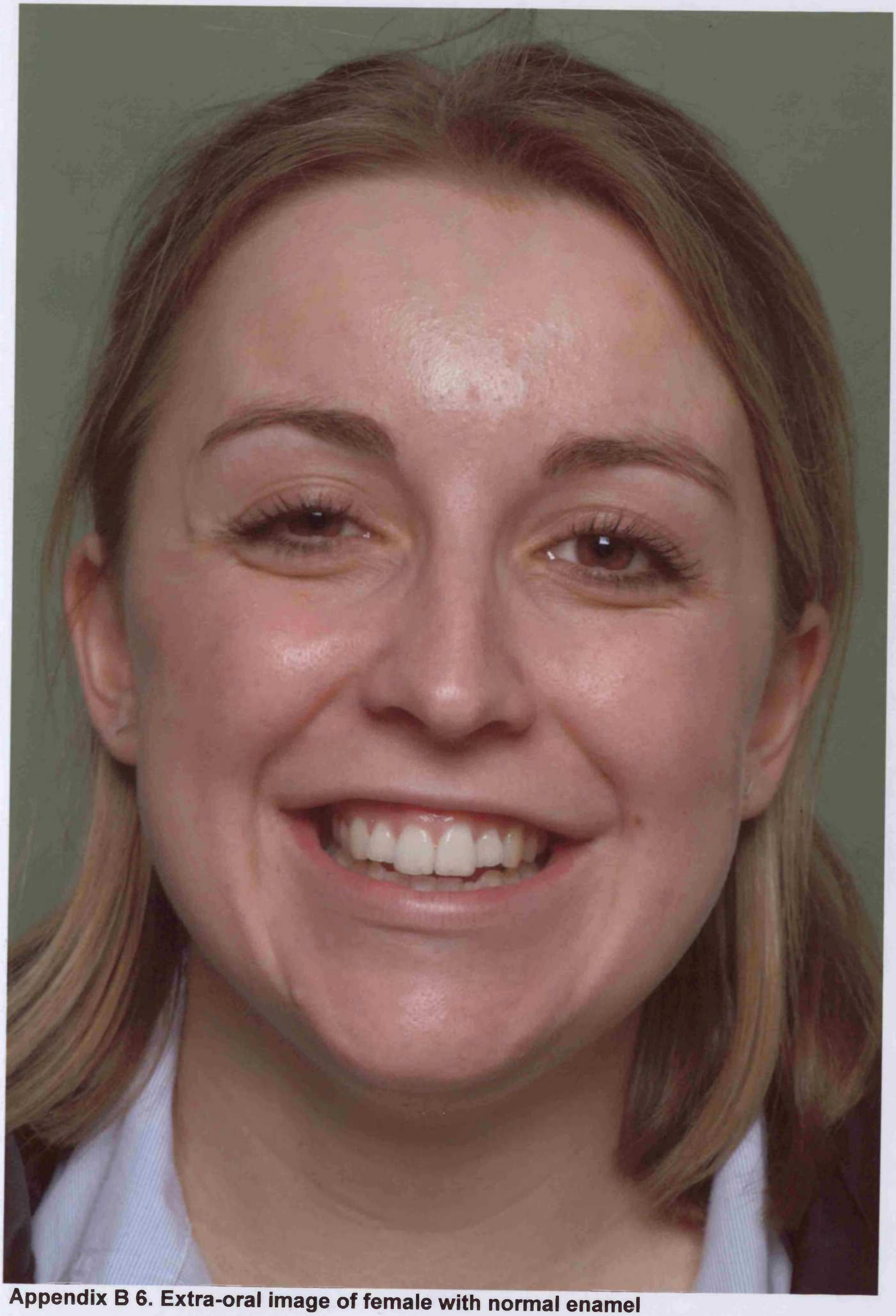
Appendix B 3. Extra-oral image of male with moderate fluorosis



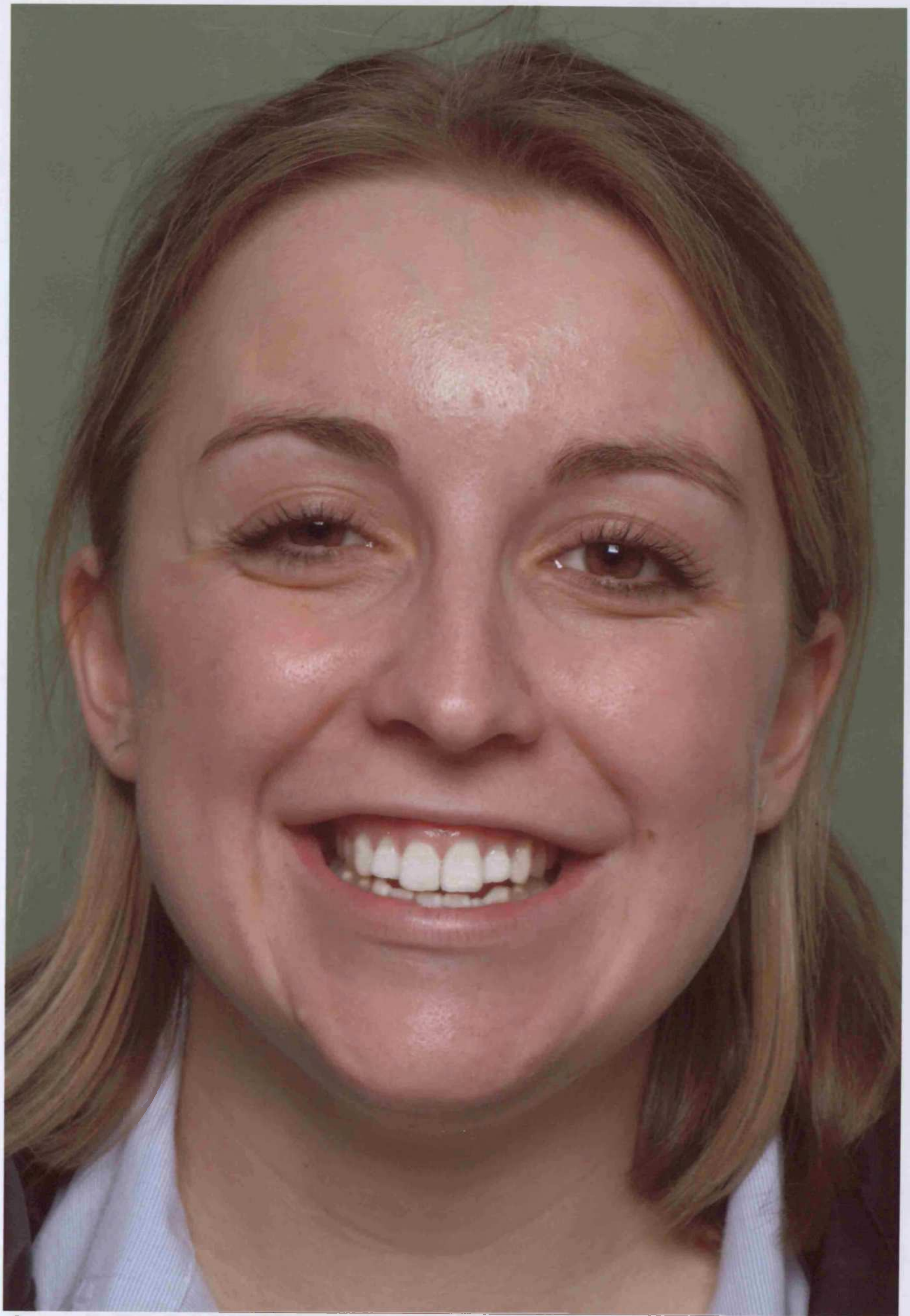
Appendix B 4. Extra-oral image of male with severe fluorosis



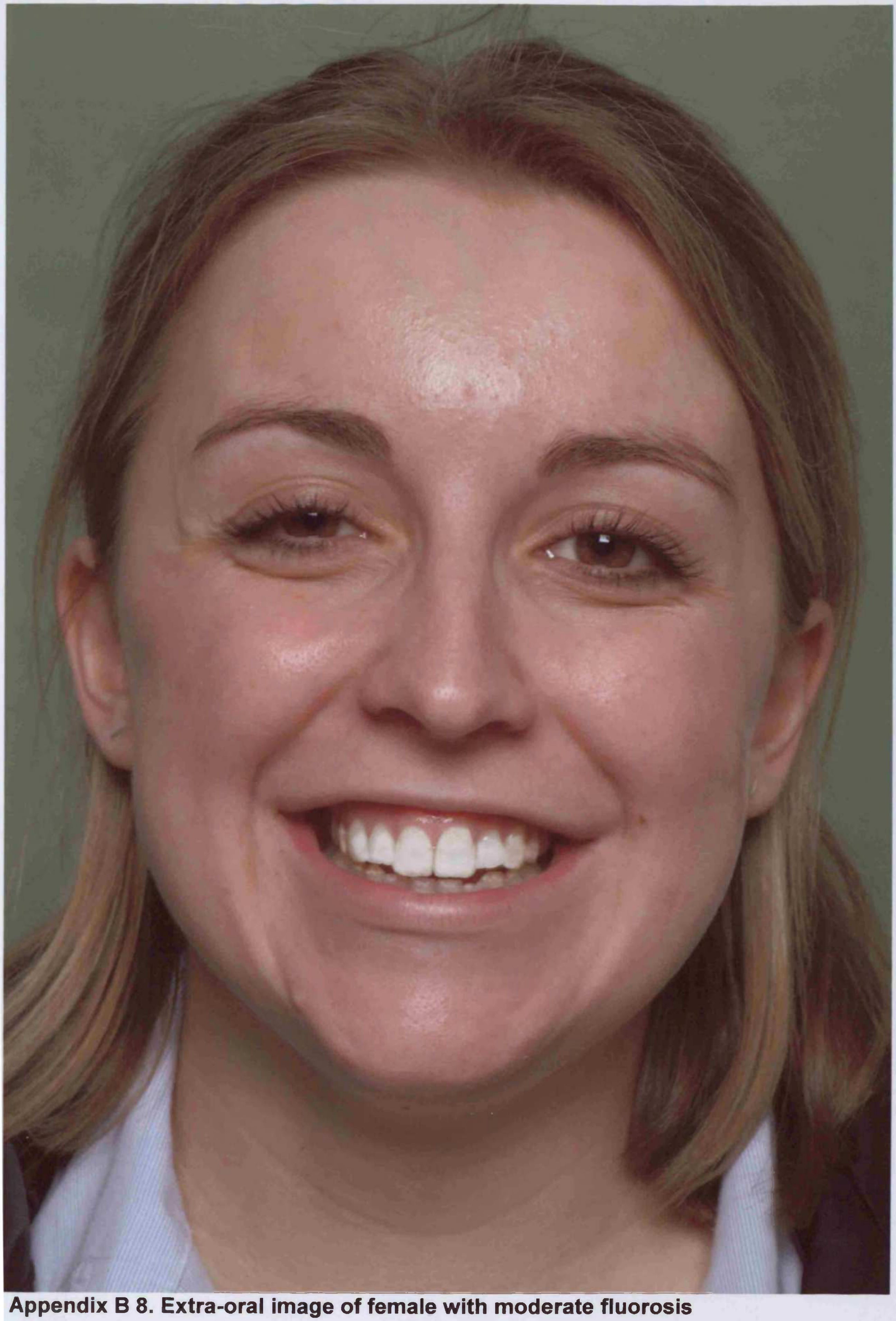
Appendix B 5. Extra-oral image of male with dental caries



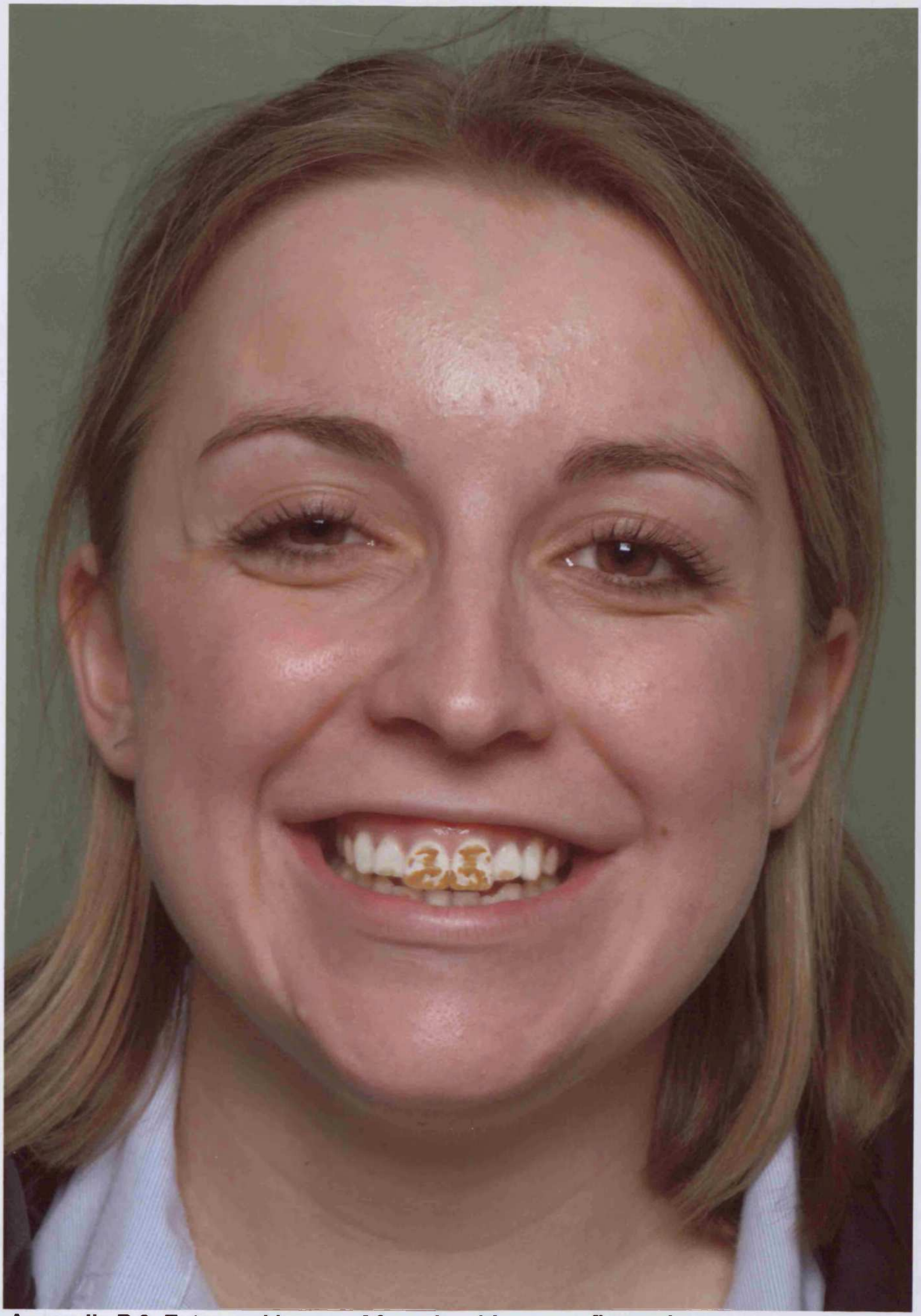
Appendix B 6. Extra-oral image of female with normal enamel



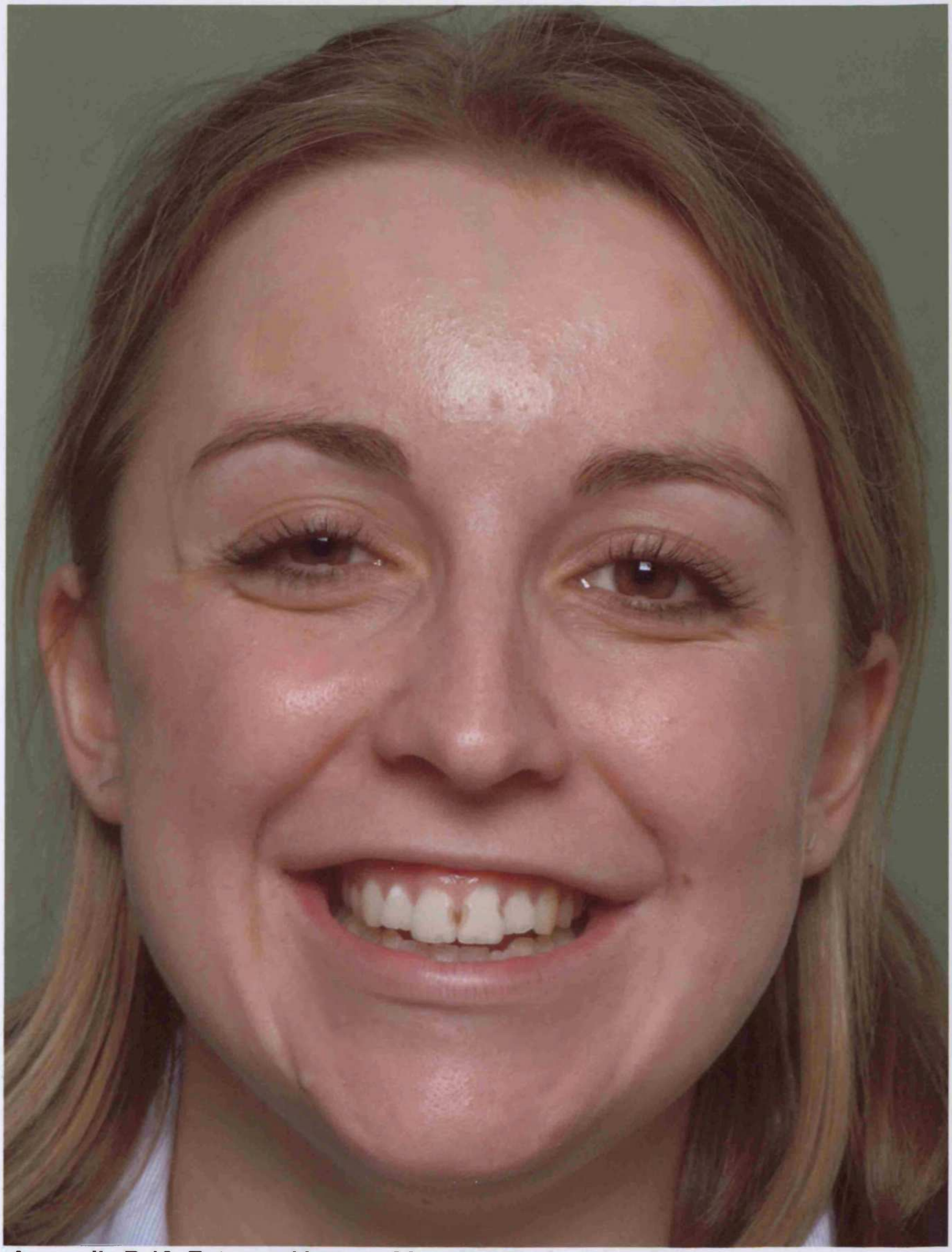
Appendix B 7. Extra-oral image of female with mild fluorosis



Appendix B 8. Extra-oral image of female with moderate fluorosis



Appendix B 9. Extra-oral image of female with severe fluorosis



Appendix B 10. Extra-oral image of female with dental caries

Appendix B 11. Intra-oral image of male with normal enamel

Appendix B 12. Intra-oral image of male with mild fluorosis

Appendix B 13. Intra-oral image of male with moderate fluorosis

Appendix B 14. Intra-oral image of male with severe fluorosis

Appendix B 15. Intra-oral image of female with normal enamel

Appendix B 16. Intra-oral image of female with mild fluorosis

Appendix B 17. Intra-oral image of female with moderate fluorosis

Appendix B 18. Intra-oral image of female with severe fluorosis

Appendix C

Materials used in the first study (Chapter 2)

- C 1. The tick list (version used with the extra-oral image without cueing to mouth).
- C 2. The tick list (version used with the extra-oral image with cueing to mouth).
- C 3. The tick list (version used with the intra-oral image).
- C 4. Recruitment leaflet.
- C 5. Information sheet (version used with the extra-oral image without cueing to mouth).
- C 6. Information sheet (version used with the extra-oral image with cueing to mouth).
- C 7. Information sheet (version used with the intra-oral image).
- C 8. Consent form.
- C 9. Demographic information sheet.

Appendix C 1

Please read the following list of traits, then tick those you think apply to the face in the photograph.

1) Beautiful ❑
2) Systematic ❑
3) Co-operative ❑
4) Sentimental ❑
5) Exciting ❑
6) Unhappy ❑
7) Tolerant ❑
8) Overcritical ❑
9) Prompt ❑
10) Unemotional ❑
11) Cowardly ❑
12) Understanding ❑
13) Frank ❑
14) Egotistical ❑
15) Rude ❑
16) Malicious ❑
17) Helpful ❑
18) Humorous ❑
19) Excited ❑
20) Stubborn ❑
21) Relaxed ❑
22) Not good looking ❑
23) Kind ❑
24) Phoney ❑
25) Sociable ❑
26) Efficient ❑
27) Intelligent ❑
28) Clean ❑
29) Unkind ❑
30) Independent ❑
31) Ugly ❑
32) Hostile ❑
33) Careless ❑
34) Gossipy ❑
35) Thoughtful ❑
36) Good looking ❑
37) Cruel ❑
38) Gullible ❑
39) Oversensitive ❑
40) Dirty ❑
41) Observant ❑
42) Unobservant ❑
43) Unpunctual ❑
44) Unhealthy ❑
45) Orderly ❑
46) Witty ❑
47) Forgiving ❑
48) Greedy ❑
49) Unconventional ❑
50) Reliable ❑

Appendix C 2

Please read the following list of traits, then tick those you think apply to the face in the photograph, please remember to pay special attention to the mouth of the face as you do this.

1) Beautiful	❑	26) Efficient	❑
2) Systematic	❑	27) Intelligent	❑
3) Co-operative	❑	28) Clean	❑
4) Sentimental	❑	29) Unkind	❑
5) Exciting	❑	30) Independent	❑
6) Unhappy	❑	31) Ugly	❑
7) Tolerant	❑	32) Hostile	❑
8) Overcritical	❑	33) Careless	❑
9) Prompt	❑	34) Gossipy	❑
10) Unemotional	❑	35) Thoughtful	❑
11) Cowardly	❑	36) Good looking	❑
12) Understanding	❑	37) Cruel	❑
13) Frank	❑	38) Gullible	❑
14) Egotistical	❑	39) Oversensitive	❑
15) Rude	❑	40) Dirty	❑
16) Malicious	❑	41) Observant	❑
17) Helpful	❑	42) Unobservant	❑
18) Humorous	❑	43) Unpunctual	❑
19) Excited	❑	44) Unhealthy	❑
20) Stubborn	❑	45) Orderly	❑
21) Relaxed	❑	46) Witty	❑
22) Not good looking	❑	47) Forgiving	❑
23) Kind	❑	48) Greedy	❑
24) Phoney	❑	49) Unconventional	❑
25) Sociable	❑	50) Reliable	❑

Appendix C 3

Please read the following list of traits, then tick those that you would associate with people who have teeth like those in the photograph.

1) Beautiful	❑	26) Efficient	❑
2) Systematic	❑	27) Intelligent	❑
3) Co-operative	❑	28) Clean	❑
4) Sentimental	❑	29) Unkind	❑
5) Exciting	❑	30) Independent	❑
6) Unhappy	❑	31) Ugly	❑
7) Tolerant	❑	32) Hostile	❑
8) Overcritical	❑	33) Careless	❑
9) Prompt	❑	34) Gossipy	❑
10) Unemotional	❑	35) Thoughtful	❑
11) Cowardly	❑	36) Good looking	❑
12) Understanding	❑	37) Cruel	❑
13) Frank	❑	38) Gullible	❑
14) Egotistical	❑	39) Oversensitive	❑
15) Rude	❑	40) Dirty	❑
16) Malicious	❑	41) Observant	❑
17) Helpful	❑	42) Unobservant	❑
18) Humorous	❑	43) Unpunctual	❑
19) Excited	❑	44) Unhealthy	❑
20) Stubborn	❑	45) Orderly	❑
21) Relaxed	❑	46) Witty	❑
22) Not good looking	❑	47) Forgiving	❑
23) Kind	❑	48) Greedy	❑
24) Phoney	❑	49) Unconventional	❑
25) Sociable	❑	50) Reliable	❑

Appendix C 4

Volunteers Needed For A Survey on How People Perceive Others Based on Their Physical Appearance.

Would you be willing to participate in a study being conducted by the University of Wales College of Medicine? The study is part of a PhD that is concerned with how people perceive others based solely on their physical appearance.

We are seeking to recruit volunteers who can spare 15 minutes of their time to help with a study investigating the public's perception of facial appearance.

What we would be asking you to do is to look at photographs of people's faces and to tell us what characteristics you think this person is likely to have, based on their appearance.

If you are willing to help with this study, we will arrange a mutually convenient time to carry out the questionnaire in the Sports and Social Club.

If you are willing to help,

Please contact David Williams on
029 2074 8275,
or by e-mail at williamsdm2@cf.ac.uk

Coleg Meddygaeth Prifysgol Cymru
University of Wales College of Medicine

Appendix C 5

An investigation of the public's perception of facial appearance

<u>Information Sheet for Participants</u>

Thank-you for considering taking part in this study. This information sheet provides details what we are proposing to do and why. Please take a few minutes to read the form.

We are carrying out a study, to find out what factors people take into account when making judgements based on facial appearance and are recruiting members of the general public to help with this project.

What we would ask you to do is to look at a series of photographs of people's faces and record any characteristics that you think this person is likely to have, based on their appearance. All information or views you express will be strictly confidential and no individuals will be specifically associated with any comments reported in the study findings.

Answering the questionnaire will take about fifteen minutes of your time and will be carried out in a quite place at the Sports and Social Club. You are under no obligation to take part. Furthermore, if having agreed to take part, you wish to withdraw at any time you are totally at liberty to do so without giving an explanation.

The study is being carried out as part of a PhD study in the University of Wales College of Medicine, Cardiff. If having helped us with this questionnaire, you are subsequently interested in finding out more about the project please contact David Williams on 029 2074 8275, or by e-mail at williamsdm2@cf.ac.uk

If you are willing to help please complete the consent form.

Pennaeth Adran / Head of Department: Yr Athro / Professor Stephen Richmond — Dental Health and Biological Sciences
incorporating Dental Public Health — Orthodontics — Paediatric Dentistry — Matrix Biology and Tissue Repair Research Unit

Ysgol Deintyddol, Mynydd Bychan, Caerdydd CF14 4XY Ffôn: +44 (0)29 2074 2447/5246/2544 Ffacs: +44 (0)29 2074 6489/4509
Dental School, Heath Park, Cardiff CF14 4XY Tel: +44 (0)29 2074 2447/5246/2544 Fax: +44 (0)29 2074 6489/4509

BUDDSODDWR MEWN POBL
INVESTOR IN PEOPLE

Appendix C 6

An investigation of the public's perception of facial appearance

Information Sheet for Participants

Thank-you for considering taking part in this study. This information sheet provides details what we are proposing to do and why. Please take a few minutes to read the form.

We are carrying out a study, to find out what factors people take into account when making judgements based on facial appearance and are recruiting members of the general public to help with this project.

What we would ask you to do is to look at a series of photographs of people's faces and record any characteristics that you think this person is likely to have, based on their appearance. Please pay particular attention to the mouth of the faces you view. All information or views you express will be strictly confidential and no individuals will be specifically associated with any comments reported in the study findings.

Answering the questionnaire will take about fifteen minutes of your time and will be carried out in a quite place at the Sports and Social Club. You are under no obligation to take part. Furthermore, if having agreed to take part, you wish to withdraw at any time you are totally at liberty to do so without giving an explanation.

The study is being carried out as part of a PhD study in the University of Wales College of Medicine, Cardiff. If having helped us with this questionnaire, you are subsequently interested in finding out more about the project please contact David Williams on 029 2074 8275, or by e-mail at williamsdm2@cf.ac.uk

If you are willing to help please complete the consent form.

Coleg Meddygaeth Prifysgol Cymru
University of Wales College of Medicine

Appendix C 7

An investigation of the public's perception of facial appearance

Information Sheet for Participants

Thank-you for considering taking part in this study. This information sheet provides details what we are proposing to do and why. Please take a few minutes to read the form.

We are carrying out a study, to find out what factors people take into account when making judgements based on facial appearance and are recruiting members of the general public to help with this project.

What we would ask you to do is to look at a series of photographs of people's teeth and record any characteristics that you think this person is likely to have, based on their dental appearance. All information or views you express will be strictly confidential and no individuals will be specifically associated with any comments reported in the study findings.

Answering the questionnaire will take about fifteen minutes of your time and will be carried out in a quite place at the Sports and Social Club. You are under no obligation to take part. Furthermore, if having agreed to take part, you wish to withdraw at any time you are totally at liberty to do so without giving an explanation.

The study is being carried out as part of a PhD study in the University of Wales College of Medicine, Cardiff. If having helped us with this questionnaire, you are subsequently interested in finding out more about the project please contact David Williams on 029 2074 8275, or by e-mail at williamsdm2@cf.ac.uk

If you are willing to help please complete the consent form.

Pennaeth Adran / Head of Department: Yr Athro / Professor Stephen Richmond — Dental Health and Biological Sciences
incorporating Dental Public Health — Orthodontics — Paediatric Dentistry — Matrix Biology and Tissue Repair Research Unit
Ysgol Deintyddol, Mynydd Bychan, Caerdydd CF14 4XY Ffôn: +44 (0)29 2074 2447/5246/2544 Ffacs: +44 (0)29 2074 6489/4509
Dental School, Heath Park, Cardiff CF14 4XY Tel: +44 (0)29 2074 2447/5246/2544 Fax: +44 (0)29 2074 6489/4509

Coleg Meddygaeth Prifysgol Cymru
University of Wales College of Medicine

Appendix C 8

Impressions of Facial Appearance

Consent Form

1. Have you read and understood the participant information sheet? **YES/NO**

2. Have you had an opportunity to discuss this study and ask any questions? **YES/NO**

3. Have you had satisfactory answers to all of your questions? **YES/NO**

4. Have you received enough information about the study? **YES/NO**

5. Who has given you an explanation about the study?

Dr./Mr/Ms..

6. Do you understand that you are free to withdraw from the study: **YES/NO**

- At any time?
- Without having to give a reason?
- That details of your participation up to the time of withdrawal will be stored anonymously on file and may be used in the final analysis of data.

7. Have you had sufficient time to come to your decision? **YES/NO**

8. Do you agree to participate in this study? **YES/NO**

PARTICIPANT
Signed...
Date...
Name (BLOCKLETTERS)..

INVESTIGATOR
Signed...
Date...
Name (BLOCKLETTERS)..

Pennaeth Adran / Head of Department: Yr Athro / Professor Stephen Richmond — Dental Health and Biological Sciences
incorporating Dental Public Health — Orthodontics — Paediatric Dentistry — Matrix Biology and Tissue Repair Research Unit
Ysgol Deintyddol, Mynydd Bychan, Caerdydd CF14 4XY Ffôn: +44 (0)29 2074 2447/5246/2544 Ffacs: +44 (0)29 2074 6489/4509
Dental School, Heath Park, Cardiff CF14 4XY Tel: +44 (0)29 2074 2447/5246/2544 Fax: +44 (0)29 2074 6489/4509

Appendix C 9

Impressions of appearance

Demographic Information

Please fill in the following details by circling the appropriate response

Gender Male
Female

Age ....

Do you work for the University of Wales College of Medicine? Yes/No

If yes, for which department do you work?

....

Appendix D

Materials used in the second and third studies (Chapters 3 and 4)

- D 1. Recruitment e-mail.
- D 2. Information sheet (version without cueing).
- D 3. Information sheet (version with cueing).
- D 4. Consent form.
- D 5. Demographic information sheet.
- D 6. On screen instructions (version without cueing).
- D 7. On screen instructions (version with cueing).
- D 8. Question sheet used to determine if participants had noticed anything unusual about the images.
- D 9. Question sheet used to assess how suitable the equipment and procedure were.
- D 10. Question sheet used to determine which dental appearances participants were explicitly aware of, and what they attributed severe fluorosis to (administered in third study only).

Appendix D 1

Would you like to earn £10 for giving us one hour of your time?

Volunteers wanted to participate in a study of first impressions based on facial appearance.

We would ask you to look at a series of images on a computer and tell us your views.

The study will last approximately one hour and will be conducted at the Student Union Building.

All volunteers will be paid £10 expenses.

If you would like to participate in this study please reply to David Williams at:
e-mail: williamsdm2@cardiff.ac.uk
telephone: 029 2074 8275.

If possible please leave your mobile phone number

Appendix D 2

An investigation of the public's perception of facial appearance

Information Sheet for Participants

Thank-you for considering taking part in this study. This information sheet provides details of what we are proposing to do and why. Please take a few minutes to read the form.

We are carrying out a study, to find out what factors people take into account when making judgements based on facial appearance and are recruiting members of the general public to help with this project.

What we would like you to do is to look at a series of images of people's faces that will be displayed on a computer monitor. After each face a word will be displayed and we would like you to tell us if you think that word applies to the previous face by pressing either the "yes" or "no" key on the response box.

All information or views you express will be strictly confidential and no individuals will be specifically associated with any comments reported in the study findings.

Completing this questionnaire will take approximately an hour of your time. You are under no obligation to take part. Furthermore, if having agreed to take part, you wish to withdraw at any time you are totally at liberty to do so without giving an explanation.

The study is being carried out as part of a PhD study in the University of Wales College of Medicine, Cardiff. If having helped us with this questionnaire, you are subsequently interested in finding out more about the project please contact David Williams on 029 2074 8275, or by e-mail at williamsdm2@cf.ac.uk

If you are willing to help please complete the consent form.

Coleg Meddygaeth Prifysgol Cymru
University of Wales College of Medicine

Appendix D 3

An investigation of the public's perception of facial appearance

Information Sheet for Participants

Thank-you for considering taking part in this study. This information sheet provides details of what we are proposing to do and why. Please take a few minutes to read the form.

We are carrying out a study, to find out what factors people take into account when making judgements based on facial appearance and are recruiting members of the general public to help with this project.

What we would like you to do is to look at a series of images of people's faces that will be displayed on a computer monitor. Please pay particular attention to the mouth of each face. After each face a word will be displayed and we would like you to tell us if you think that word applies to the previous face by pressing either the "yes" or "no" key on the response box.

All information or views you express will be strictly confidential and no individuals will be specifically associated with any comments reported in the study findings.

Completing this questionnaire will take approximately an hour of your time. You are under no obligation to take part. Furthermore, if having agreed to take part, you wish to withdraw at any time you are totally at liberty to do so without giving an explanation.

The study is being carried out as part of a PhD study in the University of Wales College of Medicine, Cardiff. If having helped us with this questionnaire, you are subsequently interested in finding out more about the project please contact David Williams on 029 2074 8275, or by e-mail at williamsdm2@cf.ac.uk

If you are willing to help please complete the consent form.

Pennaeth Adran / Head of Department: Yr Athro / Professor Stephen Richmond — Dental Health and Biological Sciences incorporating Dental Public Health — Orthodontics — Paediatric Dentistry — Matrix Biology and Tissue Repair Research Unit

Ysgol Deintyddol, Mynydd Bychan, Caerdydd CF14 4XY Ffôn: +44 (0)29 2074 2447/5246/2544 Ffacs: +44 (0)29 2074 6489/4509
Dental School, Heath Park, Cardiff CF14 4XY Tel: +44 (0)29 2074 2447/5246/2544 Fax: +44 (0)29 2074 6489/4509

Coleg Meddygaeth Prifysgol Cymru
University of Wales College of Medicine

Appendix D 4

Impressions of Facial Appearance

Consent Form

1. Have you read and understood the participant information sheet? **YES/NO**

2. Have you had an opportunity to discuss this study and ask any questions? **YES/NO**

3. Have you had satisfactory answers to all of your questions? **YES/NO**

4. Have you received enough information about the study? **YES/NO**

5. Who has given you an explanation about the study?

Dr./Mr/Ms...

6. Do you understand that you are free to withdraw from the study: **YES/NO**
- At any time?
- Without having to give a reason?
- That details of your participation up to the time of withdrawal will be stored anonymously on file and may be used in the final analysis of data.

7. Have you had sufficient time to come to your decision? **YES/NO**

8. Do you agree to participate in this study? **YES/NO**

PARTICIPANT
Signed..
Date...
Name (BLOCKLETTERS)..

INVESTIGATOR
Signed..
Date...
Name (BLOCKLETTERS)..

Pennaeth Adran / Head of Department: Yr Athro / Professor Stephen Richmond — Dental Health and Biological Sciences
incorporating Dental Public Health — Orthodontics — Paediatric Dentistry — Matrix Biology and Tissue Repair Research Unit
Ysgol Deintyddol, Mynydd Bychan, Caerdydd CF14 4XY Ffôn: +44 (0)29 2074 2447/5246/2544 Ffacs: +44 (0)29 2074 6489/4509
Dental School, Heath Park, Cardiff CF14 4XY Tel: +44 (0)29 2074 2447/5246/2544 Fax: +44 (0)29 2074 6489/4509

Appendix D 5

Impressions of appearance

Demographic Information

Please fill in the following details by circling the appropriate response

Gender Male
Female

Age ....

What course are you studying?

....

Appendix D 6

Instructions for the next stage of the experiment

In this stage of the experiment a number of faces will be displayed on the screen for a short time.

After each face a word will be displayed on the screen. Please tell us if you think that word applies to the face by pressing one of the keys on the response box.

Press the Yes key if you think the word does describe the face or the No key if you think the word does not describe the face.

After you have pressed one of the keys the next face and word will be presented. Please take care to respond accurately and as quickly as possible.

Please press either the Yes or No key when you have read these instructions and are ready to begin the next stage of the experiment.

Appendix D 7

Instructions for the next stage of the experiment

In this stage of the experiment a number of faces will be displayed on the screen for a short time.

Please pay particular attention to the mouth as you look at the face.

After each face a word will be displayed on the screen. Please tell us if you think that word applies to the face by pressing one of the keys on the response box.

Press the Yes key if you think the word does describe the face or the No key if you think the word does not describe the face.

After you have pressed one of the keys the next face and word will be presented. Please take care to respond accurately and as quickly as possible.

Please press either the Yes or No key when you have read these instructions and are ready to begin the next stage of the experiment.

Appendix D 8

An investigation of the public's perception of facial appearance – Feedback sheet

Have you noticed anything unusual about any of the images displayed in this study? If so what have you noticed?

If you have noticed anything, how confident are you that you have noticed something unusual?

Very confident	1	2	3	4	5	6	7	Not at all confident

Appendix D 9

An investigation of the public's perception of facial appearance – Evaluation sheet

How easy or difficult was it to use the response box and other equipment?

How easy or difficult was it to understand the instructions?

How easy or difficult was it to see the faces?

How easy or difficult was it to see the words?

What did you think of the choice of colour for the words and background?

What did you think about the length of time the faces were displayed for?

What did you think about the length of time that the whole study took to complete?

Any other comments?

Appendix D 10

An investigation of the public's perception of facial appearance – Feedback sheet 2

Some of the faces had brown teeth. Did you notice this? (Please circle appropriate response)

YES NO

If YES, did the image confuse you?

YES NO

Did the image surprise you?

YES NO

What did you think was the explanation for the image <u>the first time</u> you noticed it?

A computer glitch?

That the image was genuine unaltered photograph of the person?

A deliberate manipulation that was part of the study, intended to show unhealthy teeth?

A deliberate manipulation that was part of the study, intended to show staining due to food/smoking etc?

Something else (if so please tell us what)

....

....

....

Appendix E

Additional analysis and details of the second study (Chapter 3)

- E 1. Logistic regression equation.
- E 2. Linear regression equation.
- E 3. Raw descriptive data for response latencies to “yes” responses, split by characteristic.
- E 4. Raw descriptive data for response latencies to “no” responses, split by characteristic.
- E 5. Logistic regression tables split by descriptive theme.
- E 6. Linear regression tables split by descriptive theme.

Appendix E 1. The logistic regression equation

$$\left.\begin{aligned}&\text{response}_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij})\\ &\text{response}_{ij} = \pi_{ij} + e_{0ij}\text{bcons}^{*}\end{aligned}\right\}$$

$$\begin{aligned}\text{logit}(\pi_{ij}) = {}& \beta_{1j}\text{cons} + \beta_{2j}\text{valword}_{ij} + -0.042(0.104)\text{fl1}_{ij} + 0.110(0.102)\text{fl2}_{ij} + 1.756(0.095)\text{fl3}_{ij} + \\ & 0.022(0.137)\text{fl1\_val}_{ij} + -0.120(0.136)\text{fl2\_val}_{ij} + -3.264(0.129)\text{fl3\_val}_{ij}\end{aligned}$$

$$\beta_{1j} = -1.799(0.152) + u_{1j}$$

$$\beta_{2j} = 2.952(0.266) + u_{2j}$$

$$\begin{bmatrix} u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} 1.370(0.234) & \\ -2.481(0.421) & 4.767(0.790) \end{bmatrix}$$

$$\text{bcons}^{*} = \text{bcons}[\pi_{ij}(1 - \pi_{ij})/\text{denom}_{ij}]^{0.5}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim (0,\ \Omega_e) \ : \ \Omega_e = \begin{bmatrix} 0.954(0.013) \end{bmatrix}$$

Appendix E 2. The linear regression equation

Start | More | Stop | IGLS | Estimation control..

$$\text{rt\_log}_{ij} \sim \mathrm{N}(XB, \Omega)$$

$$\begin{aligned}
\text{rt_log}_{ij} = {} & \beta_{0ij}\text{cons} + 0.0014(0.0088)\text{fl1}_{ij} + 0.0105(0.0088)\text{fl2}_{ij} + 0.0843(0.0097)\text{fl3}_{ij} + \\
& 0.0950(0.0123)\text{reponse}_{ij} + 0.0627(0.0105)\text{valword}_{ij} + -0.1532(0.0165)\text{val_resp}_{ij} + \\
& 0.1043(0.0278)\text{subgen}_{j} + 0.0030(0.0094)\text{f1_val}_{ij} + 0.0095(0.0094)\text{f2_val}_{ij} + \\
& -0.0009(0.0095)\text{f3_val}_{ij} + 0.0061(0.0172)\text{f1_resp}_{ij} + 0.0173(0.0168)\text{f2_resp}_{ij} + \\
& -0.1045(0.0154)\text{f3_resp}_{ij} + 0.0042(0.0227)\text{fl1_vr}_{ij} + -0.0168(0.0225)\text{fl2_vr}_{ij} + \\
& 0.2264(0.0212)\text{fl3_vr}_{ij} + 0.0296(0.0259)\text{cueing}_{j} + -0.0094(0.0146)\text{f1_cue}_{ij} + \\
& -0.0225(0.0146)\text{f2_cue}_{ij} + -0.1237(0.0138)\text{f3_cue}_{ij}
\end{aligned}$$

$$\beta_{0ij} = 3.0151(0.0181) + u_{0j} + e_{0ij}$$

$$\left[u_{0j}\right] \sim \mathrm{N}(0, \Omega_u) : \Omega_u = \left[0.0108(0.0017)\right]$$

$$\left[e_{0ij}\right] \sim \mathrm{N}(0, \Omega_e) : \Omega_e = \left[0.0316(0.0004)\right]$$

*-2*loglikelihood(IGLS)* = -6700.7870(11340 of 11340 cases in use)

Fonts | Subs | Name | + | - | Add Term | Estimates | Nonlinear | Help | Clear

Appendix E 3. Response latencies (milliseconds), by level of fluorosis and characteristic for "yes" responses

Characteristic	N	No Fluorosis			Mild Fluorosis			Moderate Fluorosis			Severe Fluorosis		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Attractive	190	1236	1042	421-4978	1148	926	410-3966	1202	1062	480-4847	1032	876	421-2374
Unattractive	301	1763	1532	651-6259	1480	1302	771-3615	1669	1337	811-5558	1554	1276	511-7631
Careful	333	1664	1542	581-5028	1486	1262	461-4055	1450	1292	331-6549	1932	1442	380-7801
Careless	232	1539	1392	441-4577	1949	1567	711-8923	2344	1993	781-8312	1510	1362	681-4967
Clean	389	1077	942	460-3555	1296	1101	421-7040	1265	1061	451-3595	1302	922	341-5037
Dirty	193	1630	1256.5	791-4296	1842	1432	771-5118	1333	1217	401-2995	1391	1042	501-6419
Happy	532	1191	962	370-3675	1082	937	410-5187	1145	942	391-9434	1335	1152	410-5679
Unhappy	85	1505	1232	771-3545	1635	1392	451-3094	2301	2003	931-7011	1654	1312	551-6269
Healthy	392	1266	1041	510-5658	1349	1091	451-5498	1263	981	450-9804	1796	1177	411-9324
Unhealthy	220	1759	1447	731-6179	1821	1533	841-4606	2161	1633	782-7431	1546	1287	671-7831
Intelligent	396	1323	1112	520-4316	1360	1152	400-5598	1301	1046	491-6209	1629	1282	430-4897
Unintelligent	149	1798	1588	781-3375	1844	1342	881-6679	1765	1472	851-5172	1913	1597	782-8633

Appendix E 3 (continued). Response latencies (milliseconds), by level of fluorosis and characteristic for "yes" responses

Characteristic	N	No Fluorosis			Mild Fluorosis			Moderate Fluorosis			Severe Fluorosis		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Kind	524	1369	1092	431-9874	1212	1021	460-5238	1213	1031	391-4246	1639	1202	401-9474
Unkind	61	1399	1231	461-2641	1602	1422	891-2834	1623	1502	631-3224	1758	1372	641-3796
Reliable	412	1224	1102	400-4837	1318	1112	601-5478	1304	1057	520-3646	2025	1663	431-6279
Unreliable	160	2266	1658	891-8262	1938	1567	530-6460	1955	1387	942-8913	1883	1528	650-6209
Social	494	1375	1132	380-9003	1423	1142	490-6900	1358	1131	380-5998	1522	1272	500-8382
Unsocial	107	2086	1893	1051-4456	2134	1753	971-4847	1766	1748	931-2824	1892	1683	561-4437

Appendix E 4. Response latencies (milliseconds), by level of fluorosis and characteristic for "no" responses

Characteristic	N	No Fluorosis			Mild Fluorosis			Moderate Fluorosis			Severe Fluorosis		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Attractive	439	1465	1167	571-5608	1379	1166	601-5608	1303	1126	330-4016	1214	1062	390-4577
Unattractive	330	1380	1142	381-5538	1346	1172	340-4977	1623	1412	490-5398	1443	1057	411-4807
Careful	297	1925	1623	611-6710	1678	1412	341-6159	1732	1442	781-4346	1496	1267	491-4356
Careless	396	1298	1022	420-7131	1359	1162	421-6049	1587	1252	561-7521	1528	1307	431-6530
Clean	239	1542	1332	301-5989	1493	1307	420-4246	1599	1347	861-3605	1198	1092	531-4687
Dirty	437	1074	842	330-6069	1086	911	370-6660	1017	841	470-5768	1130	881	411-2894
Happy	98	1807	1572	832-5358	1495	1352	611-4467	1461	1332	671-2964	1580	1327	451-4446
Unhappy	546	1200	1012	320-4457	1239	991	430-7371	1481	1202	501-7381	1739	1362	370-8022
Healthy	240	1545	1397	311-5068	1529	1472	671-3124	1722	1412	441-5107	1347	1122	511-4436
Unhealthy	410	1271	1087	431-4246	1231	1062	511-3996	1270	1051	501-7120	1712	1242	441-7240
Intelligent	233	1683	1382	451-4046	1631	1262	341-4967	1456	1151	331-4597	1646	1397	421-6109
Unintelligent	479	1412	1162	380-8082	1396	1212	441-6029	1454	1207	541-4377	1651	1332	431-9584

Appendix E 4 (continued). Response latencies (milliseconds), by level of fluorosis and characteristic for "no" responses

Characteristic	N	No Fluorosis			Mild Fluorosis			Moderate Fluorosis			Severe Fluorosis		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Kind	108	1637	1061	601-3896	1334	1262	501-2634	1385	1362	681-2403	1418	1272	400-3655
Unkind	569	1222	1021	340-4807	1360	1142	551-8032	1257	982	401-9744	1413	1212	531-6289
Reliable	217	1324	1142	421-3334	2006	1533	530-7912	1707	1422	461-4747	1559	1241	341-7120
Unreliable	471	1459	1082	351-9934	1337	1062	380-4987	1280	1102	360-7441	1599	1287	390-9523
Social	136	1548	1372	521-3936	1565	1382	761-2673	1568	1337	751-4236	1587	1282	511-4206
Unsocial	525	1358	1052	421-6539	1281	1042	320-8713	1242	1082	341-3686	1731	1232	331-7100

Appendix E 5

Attractiveness

Table E 5.1. Two-level logistic regression model of response direction for the theme of Attractiveness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	1.18	(0.43, 3.26)	ns
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.88	(0.51, 1.51)	
Moderate	0.88	(0.51, 1.50)	
Severe	14.53	(7.99, 26.41)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.55	(0.25, 1.20)	
Moderate	0.93	(0.44, 2.00)	
Severe	0.00	(0.00, 0.01)	

[a] No fluorosis is the reference category

Carefulness

Table E 5.2. Two-level logistic regression model of response direction for the theme of Carefulness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	6.05	(2.73, 13.41)	p<0.001
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.29	(0.75, 2.22)	
Moderate	1.53	(0.90, 2.61)	
Severe	9.33	(5.45, 17.58)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.08	(0.52, 2.27)	
Moderate	0.89	(0.43, 1.87)	
Severe	0.02	(0.01, 0.05)	

[a] No fluorosis is the reference category

Cleanliness

Table E 5.3. Two-level logistic regression model of response direction for the theme of Cleanliness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	86.40	(27.13, 275.16)	p<0.001
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.35	(0.14, 0.87)	
Moderate	0.85	(0.38, 1.90)	
Severe	151.56	(62.86, 365.42)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	2.85	(0.94, 8.67)	
Moderate	0.85	(0.31, 2.35)	
Severe	0.00	(0.00, 0.00)	

[a] No fluorosis is the reference category

Happiness

Table E 5.4. Two-level logistic regression model of response direction for the theme of Happiness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	295.89	(91.11, 960.97)	p<0.001
Level of fluorosis			ns
No	1 [a]	R	
Mild	0.91	(0.45, 1.86)	
Moderate	1.74	(0.90, 3.38)	
Severe	1.75	(0.90, 3.41)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.98	(0.73, 5.36)	
Moderate	0.80	(0.31, 2.04)	
Severe	0.19	(0.08, 0.47)	

[a] No fluorosis is the reference category

Health

Table E 5.5. Two-level logistic regression model of response direction for the theme of Health

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	32.43	(12.56, 83.73)	p<0.001
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.06	(0.56, 1.99)	
Moderate	1.59	(0.86, 2.93)	
Severe	33.99	(17.77, 65.02)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.91	(0.39, 2.11)	
Moderate	0.55	(0.24, 1.27)	
Severe	0.00	(0.00, 0.00)	

[a] No fluorosis is the reference category

Intelligence

Table E 5.6. Two-level logistic regression model of response direction for the theme of Intelligence

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	22.26	(7.62, 65.05)	p<0.001
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.73	(0.40, 1.31)	
Moderate	0.86	(0.49, 1.53)	
Severe	2.74	(1.60, 4.69)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.38	(0.63, 3.00)	
Moderate	1.67	(0.77, 3.63)	
Severe	0.08	(0.04, 0.18)	

[a] No fluorosis is the reference category

Kindness

Table E 5.7. Two-level logistic regression model of response direction for the theme of Kindness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	1044.19	(272.17, 4006.05)	p<0.001
Level of fluorosis			p<0.05
No	1 [a]	R	
Mild	0.76	(0.31, 1.87)	
Moderate	0.70	(0.28, 1.75)	
Severe	2.13	(0.94, 4.81)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.97	(0.36, 2.58)	
Moderate	0.65	(0.25, 1.72)	
Severe	0.08	(0.03, 0.20)	
Cueing	0.65	(0.29, 1.46)	ns
Cueing*Level of fluorosis			p<0.05
No	1 [a]	R	
Mild	1.35	(0.52, 3.47)	
Moderate	2.78	(1.08, 7.13)	
Severe	3.03	(1.23, 7.29)	

[a] No fluorosis is the reference category

Reliability

Table E 5.8. Two-level logistic regression model of response direction for the theme of Reliability

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	45.56	(15.29, 135.74)	p<0.001
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.20	(0.60, 2.40)	
Moderate	0.64	(0.31, 1.31)	
Severe	3.92	(2.00, 7.69)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.96	(0.44, 2.10)	
Moderate	1.13	(0.51, 2.49)	
Severe	0.03	(0.01, 0.06)	
Cueing	0.95	(0.52, 1.72)	ns
Cueing*Level of fluorosis			p<0.01
No	1 [a]	R	
Mild	0.68	(0.31, 1.47)	
Moderate	1.64	(0.75, 3.58)	
Severe	2.56	(1.20, 5.44)	

[a] No fluorosis is the reference category

Sociability

Table E 5.9. Two-level logistic regression model of response direction for the theme of Sociability

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	415.30	(119.40, 1444.56)	p<0.001
Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	1.71	(0.72, 4.05)	
Moderate	1.33	(0.55, 3.22)	
Severe	10.90	(4.90, 24.26)	
Valence of Characteristic* Level of fluorosis			p<0.001
No	1 [a]	R	
Mild	0.55	(0.18, 1.72)	
Moderate	0.57	(0.18, 1.81)	
Severe	0.01	(0.00, 0.11)	

[a] No fluorosis is the reference category

Appendix E 6

Attractiveness

Table E 6.1. Two-level linear regression model showing the differences in response latency for the theme of Attractiveness

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			p<0.01
No	0 [a]	R	
Mild	-0.023	(-0.060, 0.014)	
Moderate	0.040	(0.003, 0.077)	
Severe	-0.021	(-0.060, 0.018)	
Response direction	0.059	(0.027, 0.090)	p<0.001
Valence of characteristic	-0.017	(-0.059, 0.025)	ns
Valence of characteristic* Response direction	-0.082	(-0.130, -0.033)	p<0.01
Cueing	0.071	(0.024, 0.119)	p<0.01
Cueing* Level of fluorosis			p<0.05
No	0 [a]	R	
Mild	0.019	(-0.041, 0.064)	
Moderate	-0.061	(-0.114, 0.008)	
Severe	-0.025	(-0.080, 0.029)	

[a] No fluorosis is the reference category

Carefulness

Table E 6.2. Two-level linear regression model showing the differences in response latency for the theme of Carefulness

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			ns
No	0[a]	R	
Mild	0.021	(-0.031, 0.072)	
Moderate	0.052	(0.000, 0.052)	
Severe	0.055	(-0.006, 0.116)	
Response direction	0.043	(-0.021, 0.107)	ns
Valence of characteristic	0.145	(0.094, 0.196)	p<0.001
Valence of characteristic* Response direction	-0.091	(-0.176, -0.005)	p<0.05
Level of fluorosis* Valence of characteristic			ns
No	0 [a]	R	
Mild	0.013	(-0.040, 0.066)	
Moderate	0.031	(-0.022, 0.083)	
Severe	0.017	(-0.036, 0.069)	
Level of fluorosis *Response direction			p<0.01
No	0 [a]	R	
Mild	0.055	(-0.031, 0.141)	
Moderate	0.099	(0.014, 0.184)	
Severe	-0.041	(-0.127, 0.045)	
Level of fluorosis* Valence of characteristic *Response direction			p<0.001
No	0 [a]	R	
Mild	-0.011	(-0.127, 0.104)	
Moderate	-0.107	(-0.221, -0.107)	
Severe	0.198	(0.080, 0.315)	
Cueing	0.057	(-0.008, 0.123)	ns
Cueing* Level of fluorosis			p<0.001
No	0 [a]	R	
Mild	-0.101	(0.175, -0.101)	
Moderate	-0.111	(-0.184, -0.037)	
Severe	-0.164	(-0.240, -0.089)	

[a] No fluorosis is the reference category

Cleanliness

Table E 6.3. Two-level linear regression model showing the differences in response latency for the theme of Cleanliness

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			ns
No	0[a]	R	
Mild	-0.005	(-0.041, 0.050)	
Moderate	-0.025	(-0.071, 0.021)	
Severe	0.011	(-0.061, 0.081)	
Response direction	0.127	(0.058, 0.197)	p<0.001
Valence of characteristic	0.116	(0.054, 0.179)	p<0.001
Valence of characteristic* Response direction	-0.233	(-0.330, -0.136)	p<0.001
Level of fluorosis* Valence of characteristic			ns
No	0[a]	R	
Mild	0.012	(-0.038, 0.062)	
Moderate	0.005	(-0.045, 0.054)	
Severe	0.044	(-0.006, 0.044)	
Level of fluorosis *Response direction			ns
No	0[a]	R	
Mild	0.024	(-0.086, 0.133)	
Moderate	-0.049	(-0.146, 0.049)	
Severe	-0.080	(-0.0176, 0.044)	
Level of fluorosis* Valence of characteristic *Response			ns
direction	0[a]	R	
No	0.026	(-0.115, 0.166)	
Mild	0.055	(-0.074, 0.184)	
Moderate	0.182	(0.044, 0.320)	
Severe			
Cueing	0.046	(-0.012, 0.104)	ns
Cueing* Level of fluorosis			p<0.01
No	0[a]	R	
Mild	0.005	(-0.081, 0.090)	
Moderate	0.079	(-0.005, 0.162)	
Severe	-0.096	(-0.187, -0.005)	

[a] No fluorosis is the reference category

Happiness

Table E 6.4. Two-level linear regression model showing the differences in response latency for the theme of Happiness

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			p<0.001
No	0[a]	R	
Mild	0.018	(-0.028, 0.064)	
Moderate	0.084	(0.037, 0.131)	
Severe	0.151	(0.104, 0.197)	
Response direction	0.076	(-0.007, 0.160)	ns
Valence of characteristic	0.145	(0.070, 0.219)	p<0.001
Valence of characteristic* Response direction	-0.218	(-0.332, -0.104)	p<0.001
Level of fluorosis* Valence of characteristic			ns
No	0[a]	R	
Mild	-0.017	(-0.068, 0.034)	
Moderate	-0.010	(-0.061, 0.041)	
Severe	-0.026	(-0.077, 0.025)	
Level of fluorosis *Response direction			p<0.001
No	0[a]	R	
Mild	0.036	(-0.080, 0.152)	
Moderate	0.090	(-0.018, 0.198)	
Severe	-0.121	(-0.230, -0.012)	
Level of fluorosis* Valence of characteristic *Response direction			p<0.001
No	0[a]	R	
Mild	-0.011	(-0.173, 0.152)	
Moderate	-0.074	(-0.227, 0.079)	
Severe	0.239	(0.094, 0.385)	
Cueing	0.076	(0.010, 0.143)	p<0.05
Cueing* Level of fluorosis			p<0.001
No	0[a]	R	
Mild	-0.073	(-0.186, 0.039)	
Moderate	-0.136	(-0.244, -0.0279)	
Severe	-0.208	(-0.302, -0.114)	

[a] No fluorosis is the reference category

Health

Table E 6.5. Two-level linear regression model showing the differences in response latency for the theme of Health

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			ns
No	0[a]	R	
Mild	-0.010	(-0.061, 0.040)	
Moderate	-0.022	(-0.073, 0.029)	
Severe	0.074	(0.004, 0.144)	
Response direction	0.114	(0.044, 0.184)	p<0.01
Valence of characteristic	0.059	(-0.005, 0.123)	ns
Valence of characteristic* Response direction	-0.187	(-0.285, -0.089)	p<0.001
Level of fluorosis* Valence of characteristic			ns
No	0[a]	R	
Mild	0.005	(-0.049, 0.059)	
Moderate	0.046	(-0.008, 0.100)	
Severe	0.055	(0.000, 0.110)	
Level of fluorosis *Response direction			p<0.001
No	0[a]	R	
Mild	0.027	(-0.068, 0.123)	
Moderate	0.058	(-0.035, 0.150)	
Severe	-0.150	(-0.244, -0.150)	
Level of fluorosis* Valence of characteristic *Response direction			p<0.001
No	0[a]	R	
Mild	-0.024	(-0.154, 0.106)	
Moderate	-0.077	(-0.205, 0.051)	
Severe	0.309	(0.174, 0.443)	
Cueing	0.044	(-0.018, 0.105)	ns
Cueing* Level of fluorosis			p<0.001
No	0[a]	R	
Mild	0.023	(-0.065, 0.110)	
Moderate	0.023	(-0.064, 0.110)	
Severe	-0.150	(-0.240, -0.060)	

[a] No fluorosis is the reference category

Intelligence

Table E 6.6. Two-level linear regression model showing the differences in response latency for the theme of Intelligence

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			p<0.001
No	0[a]	R	
Mild	-0.008	(-0.036, 0.020)	
Moderate	-0.014	(-0.042, 0.013)	
Severe	0.042	(0.014, 0.070)	
Response direction	0.063	(0.026, 0.100)	p<0.001
Valence of characteristic	0.014	(-0.016, 0.044)	ns
Valence of characteristic* Response direction	-0.099	(-0.153, -0.046)	p<0.001

[a] No fluorosis is the reference category

Kindness

Table E 6.7. Two-level linear regression model showing the differences in response latency for the theme of Kindness

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			p<0.001
No	0[a]	R	
Mild	0.042	(0.005, 0.079)	
Moderate	0.001	(-0.037, 0.038)	
Severe	0.068	(0.030, 0.105)	
Response direction	0.081	(0.031, 0.130)	p<0.01
Valence of characteristic	0.062	(0.012, 0.113)	p<0.05
Valence of characteristic* Response direction	-0.104	(-0.171, -0.037)	p<0.01
Cueing	0.056	(0.001, 0.112)	p<0.05
Cueing* Level of fluorosis			ns
No	0[a]	R	
Mild	-0.071	(-0.124, -0.019)	
Moderate	-0.036	(-0.080, 0.017)	
Severe	-0.022	(-0.074, 0.031)	

[a] No fluorosis is the reference category

Reliability

Table E 6.8. Two-level linear regression model showing the differences in response latency for the theme of Reliability

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			p<0.001
No	0[a]	R	
Mild	-0.019	(-0.070, 0.031)	
Moderate	-0.029	(-0.079, 0.020)	
Severe	0.099	(0.046, 0.152)	
Response direction	0.145	(0.076, 0.213)	p<0.001
Valence of characteristic	-0.212	(-0.082, 0.040)	ns
Valence of characteristic* Response direction	-0.158	(-0.253, -0.062)	p<0.01
Level of fluorosis* Valence of characteristic			ns
No	0[a]	R	
Mild	0.001	(-0.053, 0.054)	
Moderate	0.006	(-0.048, 0.059)	
Severe	-0.054	(-0.107, -0.000)	
Level of fluorosis *Response direction			p<0.01
No	0[a]	R	
Mild	0.012	(-0.083, 0.106)	
Moderate	-0.028	(-0.125, 0.068)	
Severe	-0.129	(-0.216, -0.042)	
Level of fluorosis* Valence of characteristic *Response direction			p<0.001
No	0[a]	R	
Mild	-0.102	(-0.229, 0.025)	
Moderate	-0.072	(-0.201, 0.056)	
Severe	0.311	(0.190, 0.431)	
Cueing	0.085	(0.013, 0.157)	p<0.05
Cueing* Level of fluorosis			p<0.001
No	0[a]	R	
Mild	0.147	(0.064, 0.230)	
Moderate	0.141	(0.058, 0.223)	
Severe	-0.039	(-0.117, 0.040)	

[a] No fluorosis is the reference category

Sociability

Table E 6.9. Two-level linear regression model showing the differences in response latency for the theme of Sociability

	Coefficient	95% Confidence Interval	Significance
Level of fluorosis			$p<0.001$
No	0[a]	R	
Mild	-0.013	(-0.053, 0.026)	
Moderate	-0.014	(-0.053, 0.026)	
Severe	0.075	(0.034, 0.116)	
Response direction	0.112	(0.069, 0.155)	$p<0.001$
Valence of characteristic	0.032	(-0.021, 0.084)	ns
Valence of characteristic* Response direction	-0.131	(-0.193, -0.069)	$p<0.001$
Cueing	0.064	(0.009, 0.118)	$p<0.05$
Cueing* Level of fluorosis			ns
No	0[a]	R	
Mild	0.026	(-0.030, 0.026)	
Moderate	0.013	(-0.043, 0.069)	
Severe	-0.037	(-0.094, 0.021)	

[a] No fluorosis is the reference category

Appendix F

Additional analysis and details of the third study (Chapter 4)

- F 1. Logistic regression equation.
- F 2. Linear regression equation.
- F 3. Raw descriptive data for response latencies to “yes” responses, split by characteristic.
- F 4. Raw descriptive data for response latencies to “no” responses, split by characteristic.
- F 5. Logistic regression tables split by descriptive theme.
- F 6. Linear regression tables split by descriptive theme.

Appendix F 1. The logistic regression equation

Start | More | Stop | IGLS | Estimation control..

$$\left.\begin{array}{l}\text{response}_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij}) \\ \text{response}_{ij} = \pi_{ij} + e_{0ij}\text{bcons}^*\end{array}\right\}$$

$$\text{logit}(\pi_{ij}) = \beta_{1j}\text{cons} + \beta_{2j}\text{valword}_{ij} + 0.161(0.174)\text{fl1}_{ij} + 1.594(0.156)\text{caries}_{ij} + 2.200(0.155)\text{fl3}_{ij} + -0.411(0.233)\text{fl1\_val}_{ij} + -3.470(0.213)\text{car\_val}_{ij} + -4.698(0.215)\text{fl3\_val}_{ij}$$

$$\beta_{1j} = -2.409(0.208) + u_{1j}$$

$$\beta_{2j} = 4.379(0.357) + u_{2j}$$

$$\begin{bmatrix} u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0, \Omega_u) : \Omega_u = \begin{bmatrix} 1.050(0.257) & \\ -1.923(0.467) & 3.847(0.902) \end{bmatrix}$$

$$\text{bcons}^* = \text{bcons}[\pi_{ij}(1 - \pi_{ij})/\text{denom}_{ij}]^{0.5}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim (0, \Omega_e) : \Omega_e = \begin{bmatrix} 1.032(0.019) \end{bmatrix}$$

Fonts | Subs | Name | + | - | Add Term | Estimates | Nonlinear | Help | Clear

Appendix F 2. The linear regression equation

Start | More | Stop | IGLS | Estimation control..

$$\text{rt\_log}_{ij} \sim \mathrm{N}(XB, \Omega)$$

$$\begin{aligned}\text{rt\_log}_{ij} = {} & \beta_{0ij}\text{cons} + 0.0071(0.0108)\text{fl1}_{ij} + 0.0364(0.0116)\text{caries}_{ij} + 0.0733(0.0123)\text{fl3}_{ij} + \\ & 0.1365(0.0225)\text{response}_{ij} + 0.0579(0.0194)\text{valword}_{ij} + -0.2333(0.0300)\text{res\_val}_{ij} + \\ & -0.0244(0.0308)\text{fl1\_res}_{ij} + -0.1279(0.0271)\text{car\_res}_{ij} + -0.1427(0.0267)\text{fl3\_res}_{ij} + \\ & -0.0084(0.0262)\text{fl1\_val}_{ij} + -0.0483(0.0236)\text{car\_val}_{ij} + -0.1111(0.0235)\text{fl3\_val}_{ij} + \\ & 0.0379(0.0406)\text{fl1\_vr}_{ij} + 0.2288(0.0364)\text{car\_vr}_{ij} + 0.2766(0.0363)\text{fl3\_vr}_{ij}\end{aligned}$$

$$\beta_{0ij} = 3.0783(0.0203) + u_{0j} + e_{0ij}$$

$$\left[u_{0j}\right] \sim \mathrm{N}(0, \Omega_u) : \Omega_u = \left[0.0142(0.0032)\right]$$

$$\left[e_{0ij}\right] \sim \mathrm{N}(0, \Omega_e) : \Omega_e = \left[0.0364(0.0007)\right]$$

*-2*loglikelihood(IGLS)* = **-2561.9810(5746 of 5746 cases in use)**

Fonts | Subs | Name | + | - | Add Term | Estimates | Nonlinear | Help | Clear

Appendix F 3. Response latencies (milliseconds), by dental appearance and characteristic for "yes" responses

Characteristic	N	Normal enamel			Mild fluorosis			Severe fluorosis			Dental caries		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Attractive	134	1153	1002	450-3155	1137	891	491-3836	3249	2013	551-9884	1737	1292	621-4877
Unattractive	153	2332	1993	891-5097	1972	1588	801-6769	1552	1282	721-5969	1501	1282	781-3916
Careful	173	1469	1141	461-7131	1732	1262	391-6259	3137	1252	691-2133	2074	1592	561-9403
Careless	117	1972	1442	350-6710	1511	1472	561-3024	1656	1262	711-7901	1658	1202	711-4867
Clean	175	1120	881	341-5308	1214	972	391-3826	1476	1412	621-3645	1700	1072	581-6009
Dirty	123	1934	1392	841-4156	1592	1422	831-2704	1225	1107	531-4647	1200	1052	420-3716
Happy	258	1080	836	351-4176	1004	926	451-2514	1349	1102	531-4576	1307	1172	501-2774
Unhappy	38	1798	1077	881-5648	1729	1722	1472-1993	1894	1793	931-4116	1884	1447	821-4447
Healthy	191	1154	956	441-5538	1212	931	430-5668	1599	1382	581-4917	1840	1462	551-9904
Unhealthy	110	1806	1802	771-3025	2116	1893	1091-3846	1523	1171	651-5428	1436	1251	791-3295
Intelligent	202	1219	982	451-3695	1237	1052	510-4737	1576	1403	581-4356	1449	1242	601-4226
Unintelligent	66	2418	2153	1062-5558	1781	1472	901-3175	1866	1783	821-4887	1751	1292	831-5929

Appendix F 3 (continued). Response latencies (milliseconds), by dental appearance and characteristic for "yes" responses

Characteristic	N	Normal enamel			Mild fluorosis			Severe fluorosis			Dental caries		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Kind	267	1430	1006	451-7631	1278	992	430-3595	1739	1443	490-4296	1387	1197	541-4537
Unkind	25	2634	3074	621-5318	2178	1948	591-4466	2789	2013	630-8682	1664	1572	811-2874
Reliable	211	1479	1151	381-4466	1565	1221	370-4897	1968	1592	601-5378	1905	1412	610-8192
Unreliable	67	1854	1862	1052-2623	2104	1498	862-4296	1672	1492	881-3305	1562	1442	801-4186
Social	235	1403	1021	371-7160	1380	957	350-5418	1692	1207	621-8372	1592	1277	621-8472
Unsocial	49	1552	1552	1552-1552	1818	1708	1553-2404	1728	1372	781-5829	1675	1167	461-4016

Appendix F 4. Response latencies (milliseconds), by dental appearance and characteristic for “no” responses

Characteristic	N	Normal enamel			Mild fluorosis			Severe fluorosis			Dental caries		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Attractive	186	1479	1352	721-3555	1229	1032	661-3034	1298	1051	651-3755	1391	1202	631-4887
Unattractive	167	1437	1291	441-5248	1654	1442	511-5618	1187	901	501-3445	1392	1122	550-5488
Careful	147	1787	1332	550-6570	1521	1302	541-4336	1428	1072	520-4927	1768	1337	701-8102
Careless	201	1535	1257	570-7241	1305	1152	511-4597	1580	1357	430-5278	1784	1322	651-8141
Clean	144	1434	1513	761-2394	1697	1131	681-5497	1143	931	551-3806	1206	1071	591-2894
Dirty	196	1136	992	451-3695	1174	981	451-3144	2406	1528	561-8823	1721	1312	511-6789
Happy	62	1542	1352	611-3295	1544	1392	521-3455	1488	1121	631-5408	2097	1352	561-6539
Unhappy	281	1177	921	430-5258	1489	1122	491-5829	1407	1091	590-5818	1411	1112	541-4567
Healthy	128	1609	1187	551-4897	1511	1332	661-3125	1153	1052	621-2995	1388	1026	681-3756
Unhealthy	210	1318	1131	381-4437	1596	1282	521-7561	1807	1492	450-4106	1661	1162	541-8232
Intelligent	118	1319	1232	521-2193	1767	1412	711-4106	1507	1212	530-4406	1938	1392	681-9244
Unintelligent	252	1335	1062	530-4697	1352	1042	541-4426	1829	1362	580-6479	1363	1121	501-4457

Appendix F 4 (continued). Response latencies (milliseconds), by dental appearance and characteristic for "no" responses

Characteristic	N	Normal enamel			Mild fluorosis			Severe fluorosis			Dental caries		
		Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
Kind	53	2530	2048	801-4997	2554	1582	621-7361	1756	1172	771-4537	1665	1067	491-5338
Unkind	293	1401	1092	350-5979	1299	982	471-5067	1607	1453	721-6990	1352	1191	501-3685
Reliable	108	1438	1567	581-2073	1727	1281	801-7020	1665	1472	671-5848	1459	1171	601-3875
Unreliable	250	1514	1322	521-4396	1408	1237	520-4697	1848	1327	561-8792	1773	1642	631-7962
Social	85	2259	1251	621-5117	1879	1878	821-3785	1552	1513	551-3786	1449	1212	500-3856
Unsocial	271	1395	1131	520-4346	1334	1097	531-4697	1846	1247	521-8743	1456	1192	490-6429

Appendix F 5

Attractiveness

Table F 5.1. Two-level logistic regression model of response direction for the theme of Attractiveness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	10.24	(2.99, 35.10)	p<0.001
Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	2.45	(1.06, 5.66)	
Dental caries	11.48	(4.96, 26.56)	
Severe fluorosis	17.83	(7.55, 42.10)	
Valence of Characteristic* Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	0.47	(0.15, 1.50)	
Dental caries	0.02	(0.00, 0.05)	
Severe fluorosis	0.00	(0.00, 0.01)	

[a] Normal enamel is reference category

Carefulness

Table F 5.2. Two-level logistic regression model of response direction for the theme of Carefulness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	24.76	(8.95, 68.53)	p<0.001
Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	0.83	(0.37, 1.88)	
Dental caries	5.42	(2.59, 11.36)	
Severe fluorosis	6.57	(3.13, 13.82)	
Valence of Characteristic* Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	1.02	(0.33, 3.11)	
Dental caries	0.02	(0.01, 0.06)	
Severe fluorosis	0.01	(0.00, 0.02)	

[a] Normal enamel is reference category

Cleanliness

Table F 5.3. Two-level logistic regression model of response direction for the theme of Cleanliness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	112.87	(27.95, 455.83)	p<0.001
Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	0.53	(0.17, 1.64)	
Dental caries	22.01	(8.52, 56.87)	
Severe fluorosis	52.95	(19.40, 144.49)	
Valence of Characteristic* Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	1.88	(0.44, 8.01)	
Dental caries	0.00	(0.00, 0.01)	
Severe fluorosis	0.00	(0.00, 0.00)	

[a] Normal enamel is reference category

Happiness

Table F 5.4. Two-level logistic regression model of response direction for the theme of Happiness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	1154.01	(169.81, 7842.34)	p<0.001
Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	0.39	(0.09, 1.79)	
Dental caries	2.06	(0.64, 6.63)	
Severe fluorosis	6.49	(2.16, 19.57)	
Valence of Characteristic* Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	1.79	(0.28, 2.45)	
Dental caries	0.11	(0.02, 0.50)	
Severe fluorosis	0.02	(0.00, 0.08)	

[a] Normal enamel is the reference category

Health

Table F 5.5. Two-level logistic regression model of response direction for the theme of Health

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	3.25	(1.81, 5.82)	p<0.001
Dental appearance			ns
Normal enamel	1 [a]	R	
Mild fluorosis	0.97	(0.61, 1.54)	
Dental caries	0.71	(0.44, 1.13)	
Severe fluorosis	1.20	(0.53, 1.33)	

[a] Normal enamel is the reference category

Intelligence

Table F 5.6. Two-level logistic regression model of response direction for the theme of Intelligence

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	351.18	(73.18, 1685.29)	p<0.001
Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	1.38	(0.54, 3.49)	
Dental caries	2.96	(1.22, 7.20)	
Severe fluorosis	9.73	(4.02, 23.54)	
Valence of Characteristic* Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	0.21	(0.06, 0.75)	
Dental caries	0.02	(0.01, 0.08)	
Severe fluorosis	0.00	(0.00, 0.01)	

[a] Normal enamel is the reference category

Kindness

Table F 5.7. Two-level logistic regression model of response direction for the theme of Kindness

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	658.06	(128.99, 3357.14)	p<0.001
Dental appearance			p<0.05
Normal enamel	1 [a]	R	
Mild fluorosis	0.84	(0.40, 1.77)	
Dental caries	0.43	(0.21, 0.89)	
Severe fluorosis	0.37	(0.18, 0.77)	

[a] Normal enamel is the reference category

Reliability

Table F 5.8. Two-level logistic regression model of response direction for the theme of Reliability

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	332.89	(75.17, 1474.14)	p<0.001
Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	1.18	(0.40, 3.49)	
Dental caries	3.94	(1.52, 10.21)	
Severe fluorosis	11.09	(4.39, 28.06)	
Valence of Characteristic* Dental appearance			p<0.001
Normal enamel	1 [a]	R	
Mild fluorosis	0.34	(0.08, 1.44)	
Dental caries	0.02	(0.01, 0.09)	
Severe fluorosis	0.00	(0.00, 0.01)	

[a] Normal enamel is the reference category

Sociability

Table F 5.9. Two-level logistic regression model of response direction for the theme of Sociability

	Odds Ratio	95% Confidence Interval	Significance
Valence of characteristic	2284.27	(97.26, 53647.17)	$p<0.001$
Dental appearance			$p<0.001$
Normal enamel	1 [a]	R	
Mild fluorosis	8.43	(0.37, 191.87)	
Dental caries	30.05	(1.49, 606.25)	
Severe fluorosis	116.87	(6.03, 2264.72)	
Valence of Characteristic* Dental appearance			$p<0.001$
Normal enamel	1 [a]	R	
Mild fluorosis	0.13	(0.00, 3.91)	
Dental caries	0.01	(0.00, 0.25)	
Severe fluorosis	0.00	(0.00, 0.03)	

[a] Normal enamel is the reference category

Appendix F 6.

Attractiveness

Table F 6.1. Two-level linear regression model showing the differences in response latency for the theme of Attractiveness

	Coefficient	95% Confidence Interval	Significance
Dental appearance			$p<0.05$
Normal enamel	0 [a]	R	
Mild fluorosis	0.048	(-0.017, 0.112)	
Dental caries	-0.041	(-0.118, 0.037)	
Severe fluorosis	-0.075	(-0.160, 0.009)	
Response direction	0.190	(0.094, 0.287)	$p<0.001$
Valence of characteristic	0.037	(-0.040, 0.113)	ns
Valence of characteristic* Response direction	-0.322	(-0.451, -0.193)	$p<0.001$
Dental appearance* Valence of characteristic			$p<0.05$
Normal enamel	0 [a]	R	
Mild fluorosis	-0.133	(-0.242, -0.024)	
Dental caries	0.011	(-0.099, 0.120)	
Severe fluorosis	0.015	(-0.098, 0.127)	
Dental appearance *Response direction			ns
Normal enamel	0 [a]	R	
Mild fluorosis	-0.091	(-0.215, 0.034)	
Dental caries	-0.103	(-0.229, 0.022)	
Severe fluorosis	-0.083	(-0.212, 0.046)	
Dental appearance* Valence of characteristic *Response direction			$p<0.01$
Normal enamel	0 [a]	R	
Mild fluorosis	0.176	(0.008, 0.345)	
Dental caries	0.273	(0.101, 0.446)	
Severe fluorosis	0.355	(0.158, 0.552)	

[a] Normal enamel is the reference category

Carefulness

Table F 6.2. Two-level linear regression model showing the differences in response latency for the theme of Carefulness

	Coefficient	95% Confidence Interval	Significance
Dental appearance			p<0.05
Normal enamel	0 [a]	R	
Mild fluorosis	-0.004	(-0.048, 0.039)	
Dental caries	0.046	(0.000, 0.091)	
Severe fluorosis	-0.016	(-0.062, 0.030)	
Response direction	0.056	(0.007 0.106,)	p<0.05
Valence of characteristic	0.037	(-0.009, 0.082)	ns
Valence of characteristic* Response direction	-0.077	(-0.151, -0.003)	p<0.05

[a] Normal enamel is the reference category

Cleanliness

Table F 6.3. Two-level linear regression model showing the differences in response latency for the theme of Cleanliness

	Coefficient	95% Confidence Interval	Significance
Dental appearance			p<0.001
Normal enamel	0 [a]	R	
Mild fluorosis	0.012	(-0.046, 0.070)	
Dental caries	0.125	(0.050, 0.201)	
Severe fluorosis	0.250	(0.163, 0.337)	
Response direction	0.183	(0.067, 0.298)	p<0.01
Valence of characteristic	0.078	(-0.029, 0.185)	ns
Valence of characteristic* Response direction	-0.284	(-0.444, -0.124)	p<0.001
Dental appearance* Valence of characteristic			p<0.001
Normal enamel	0 [a]	R	
Mild fluorosis	0.014	(-0.135, 0.163)	
Dental caries	-0.179	(-0.312, -0.047)	
Severe fluorosis	-0.328	(-0.466, -0.189)	
Dental appearance *Response direction			p<0.001
Normal enamel	0 [a]	R	
Mild fluorosis	-0.089	(-0.268, 0.090)	
Dental caries	-0.278	(-0.419, -0.137)	
Severe fluorosis	-0.402	(-0.549, -0.255)	
Dental appearance* Valence of characteristic *Response direction			p<0.001
Normal enamel	0 [a]	R	
Mild fluorosis	0.102	(-0.132, 0.335)	
Dental caries	0.458	(0.260, 0.656)	
Severe fluorosis	0.581	(0.368, 0.794)	

[a] Normal enamel is the reference category

Happiness

Table F 6.4. Two-level linear regression model showing the differences in response latency for the theme of Happiness

	Coefficient	95% Confidence Interval	Significance
Dental appearance			p<0.001
Normal enamel	0 [a]	R	
Mild fluorosis	0.038	(-0.001, 0.077)	
Dental caries	0.091	(0.052, 0.130)	
Severe fluorosis	0.083	(0.043, 0.123)	
Response direction	0.055	(-0.011, 0.122)	ns
Valence of characteristic	0.007	(-0.047, 0.061)	ns
Valence of characteristic* Response direction	-0.107	(-0.200, -0.014)	p<0.05

[a] Normal enamel is the reference category

Health

Table F 6.5. Two-level linear regression model showing the differences in response latency for the theme of Health

	Coefficient	95% Confidence Interval	Significance
Dental appearance			$p<0.05$
Normal enamel	0 [a]	R	
Mild fluorosis	0.051	(-0.009, 0.112)	
Dental caries	0.050	(-0.020, 0.119)	
Severe fluorosis	0.121	(0.038, 0.203)	
Response direction	0.106	(-0.019, 0.231)	ns
Valence of characteristic	0.050	(-0.075, 0.175)	ns
Valence of characteristic* Response direction	-0.229	(-0.409, -0.049)	$p<0.05$
Dental appearance* Valence of characteristic			$p<0.01$
Normal enamel	0 [a]	R	
Mild fluorosis	-0.012	(-0.183, 0.159)	
Dental caries	-0.066	(-0.212, 0.080)	
Severe fluorosis	-0.212	(-0.362, -0.061)	
Dental appearance *Response direction			$p<0.05$
Normal enamel	0 [a]	R	
Mild fluorosis	0.065	(-0.110, 0.239)	
Dental caries	-0.099	(-0.248, 0.050)	
Severe fluorosis	-0.172	(-0.324, -0.020)	
Dental appearance* Valence of characteristic *Response direction			$p<0.001$
Normal enamel	0 [a]	R	
Mild fluorosis	-0.081	(-0.328, 0.166)	
Dental caries	0.255	(0.040, 0.469)	
Severe fluorosis	0.409	(0.189, 0.630)	

[a] Normal enamel is the reference category

Intelligence

Table F 6.6. Two-level linear regression model showing the differences in response latency for the theme of Intelligence

	Coefficient	95% Confidence Interval	Significance
Dental appearance			$p<0.01$
Normal enamel	0 [a]	R	
Mild fluorosis	-0.007	(-0.066, 0.053)	
Dental caries	0.005	(-0.056, 0.066)	
Severe fluorosis	0.108	(0.044, 0.173)	
Response direction	0.159	(0.031, 0.287)	$p<0.05$
Valence of characteristic	-0.004	(-0.117, 0.110)	ns
Valence of characteristic* Response direction	-0.203	(-0.377, -0.028)	$p<0.05$
Dental appearance* Valence of characteristic			$p<0.01$
Normal enamel	0 [a]	R	
Mild fluorosis	0.103	(-0.038, 0.244)	
Dental caries	0.106	(-0.027, 0.239)	
Severe fluorosis	-0.651	(-0.199, 0.069)	
Dental appearance *Response direction			ns
Normal enamel	0 [a]	R	
Mild fluorosis	-0.070	(-0.240, 0.101)	
Dental caries	-0.103	(-0.262, 0.056)	
Severe fluorosis	-0.135	(-0.286, 0.017)	
Dental appearance* Valence of characteristic *Response direction			ns
Normal enamel	0 [a]	R	
Mild fluorosis	-0.022	(-0.244, 0.200)	
Dental caries	0.082	(-0.129, 0.293)	
Severe fluorosis	0.198	(-0.010, 0.405)	

[a] Normal enamel is the reference category

Kindness

Table F 6.7. Two-level linear regression model showing the differences in response latency for the theme of Kindness

	Coefficient	95% Confidence Interval	Significance
Dental appearance			$p<0.001$
Normal enamel	0 [a]	R	
Mild fluorosis	-0.021	(-0.062, 0.020)	
Dental caries	0.001	(-0.040, 0.042)	
Severe fluorosis	0.088	(0.047, 0.130)	
Response direction	0.141	(0.057, 0.225)	$p<0.001$
Valence of characteristic	0.067	(0.009, 0.126)	$p<0.05$
Valence of characteristic* Response direction	-0.196	(-0.303, -0.088)	$p<0.001$

[a] Normal enamel is the reference category

Reliability

Table F 6.8. Two-level linear regression model showing the differences in response latency for the theme of Reliability

	Coefficient	95% Confidence Interval	Significance
Dental appearance			$p<0.05$
Normal enamel	0 [a]	R	
Mild fluorosis	0.002	(-0.040, 0.043)	
Dental caries	0.040	(-0.003, 0.082)	
Severe fluorosis	0.061	(0.017, 0.106)	
Response direction	0.043	(-0.013, 0.099)	ns
Valence of characteristic	0.003	(-0.044, 0.050)	ns
Valence of characteristic* Response direction	-0.053	(-0.133, 0.028)	ns

[a] Normal enamel is the reference category

Sociability

Table F 6.9. Two-level linear regression model showing the differences in response latency for the theme of Sociability

	Coefficient	95% Confidence Interval	Significance
Dental appearance			ns
Normal enamel	0 [a]	R	
Mild fluorosis	-0.008	(-0.051, 0.035)	
Dental caries	0.015	(-0.028, 0.059)	
Severe fluorosis	0.047	(0.001, 0.092)	
Response direction	0.029	(-0.039, 0.096)	ns
Valence of characteristic	0.037	(-0.016, 0.090)	ns
Valence of characteristic* Response direction	-0.077	(-0.171, 0.017)	ns

[a] Normal enamel is the reference category