Evidence for Excellence in Education

Appendices

Research into the impact of Bikeability training on children's ability to perceive and appropriately respond to hazards when cycling on the road

National Foundation for Educational Research (NFER)

Research into the impact of Bikeability training on children's ability to perceive and appropriately respond to hazards when cycling on the road: Appendices

Claire Hodgson
Jack Worth
Michael Frearson

Published in January 2015
By the National Foundation for Educational Research,
The Mere, Upton Park, Slough, Berkshire SL1 2DQ
www.nfer.ac.uk

ISBN 978-1-91000846-1

How to cite this publication: Hodgson, C., Frearson, M. and Worth, J. (2015). *Research into the impact of Bikeability training on children's ability to perceive and appropriately respond to hazards when cycling on the road: Appendices*. Slough: NFER.

Contents

Project team

Catherine Kirkup	Researchers
Claire Hodgson	
Michael Frearson	
Nicola Keogh	Project Co-ordinator
Jack Worth	Statistician
Ian Brownlie	Research and Product Operations

Acknowledgements

This study would not have been possible without the active participation of schools, pupils and Bikeability schemes. We would like to thank the all participants for their valuable contributions.

We are also grateful to Professor John Wann (Professor of Psychology at Royal Holloway, University of London and a British Cycling Level 3 Coach) for: advising on the relevant accident data and literature to review, advising on the content of the on-screen quiz and reviewing the interpretation of findings.

Appendix A – Review and analysis of relevant accident data and literature relating to cycle training

A.1 Purpose

The purpose of this review is to consider the most common risks children face when cycling on the road in order to inform the design of the on-screen quiz and practical assessments for testing children's ability to perceive and appropriately respond to hazards. The review first considers common risks faced by children riding on the road and the factors which affect their ability to develop skills of hazard perception. It then identifies common on-road conflicts including those involving child cyclists; factors contributing to these conflicts; and the effectiveness of cycle training in reducing risk for children.

A.2 Introduction

In a recent report, the OECD International Transport Forum (2013) identified the many appeals of cycling - bicycles "use no fossil energy, deliver important health benefits … and provide an affordable and seamless door-to-door mobility service" (p. 37). Learning to ride a bicycle is, as Klin *et al*. (2009) point out, a "developmental milestone in the life of a child, a source of independence and freedom" (p. 1011).

However, it is also the case that cyclists are vulnerable road users and "crash outcomes are especially severe for … cyclists … single bicycle crashes are also a source of injuries through falls and collisions with obstacles and can result in serious injuries". It is also, perhaps, unsurprising that "studies investigating the comparative risk of injury for cyclists versus car occupants find significantly higher risks per unit of exposure for cyclists" (OECD International Transport Forum, 2013, p. 17). Further, in the UK, the risk of a cyclist being killed or seriously injured has been reported to be highest for young cyclists aged 10-15 years (Knowles *et al*., 2009a, p. 12).

Hazard perception is a key skill for cyclists because "a bicycle is a vehicle … (it) moves at a certain speed, which means that the cyclist needs time to react and brake or change direction if an obstacle or an unpredicted situation occurs" (OECD International Transport Forum, 2013, p. 175).

But what exactly are the risks and hazards facing cyclists, particularly children, and how might training help to reduce these risks?

A.3 Common risks faced by children and the factors affecting their ability to develop hazard perception skills

In Great Britain in 2012, the number of seriously injured pedal cyclists increased by four per cent, the eighth consecutive annual increase, while the total number of child casualties (aged 0-15) decreased by 11 per cent from 2011, the lowest total since records began in 1979 (Department for Transport, 2013).

Turner *et al.* (2009) tell us that "Depending on their age, children have serious knowledge, perceptual and cognitive limitations in relation to roads. They can be unpredictable, do not have a good appreciation of road hazards and are generally unfamiliar with road rules" (p. 17). Children aged 10-15 remain more at risk of injury than other age group in the population, and within this group the risk of injury increases with age. This may be because children generally have less awareness of their own abilities and of road priority rules than adult cyclists, most of whom also drive. Until the age of 10, most children underestimate the speed of cars, but between the ages of 10 and 14 boys (who are overrepresented in child cycling casualty figures) overestimate the speed of approaching vehicles. Moreover, the transition to secondary school often involves more independent and unsupervised travel for older children in this age group (Knowles *et al*., 2009a).

Children's ability to co-ordinate self motion with the motions of other objects appears, according to Plumert *et al*. (2011), "to undergo developmental change up until at least 12 years of age" (p. 1245). They further explain that younger children, in particular, have difficulty determining how long it will take to start up and cycle a particular distance, particularly from a dead stop, which can pose problems when attempting to judge a suitable gap in which to cycle across a road. This research is consistent with other findings that errors in judging affordances may play an important role in unintentional childhood injuries (pp. 1249–50).

However, by the age of 10, children can achieve basic cycling competence with appropriate training, for riding on quiet two-lane roads, negotiating parked cars and simple junctions; however, they may lack confidence to defend single lanes in narrow roads (Turner *et al*., 2009).

Age and experience are the main predictors of children's performance of on-road cycling skills, such as riding one-handed and looking behind while cycling, between ages five and 13 (Maring and van Schagen, 1990). Unsurprisingly, the least experienced cyclists, those riding for five years or less, are more likely to be injured than more experienced cyclists (Heesch *et al*., 2011).

A.4 Common on-road conflicts including those involving children on bicycles

Knowles *et al*. (2009a) reported that over four-fifths of cyclists killed or seriously injured (KSI) were as a result of an impact with another vehicle and that over two-thirds of these involved a collision with a car/taxi. Two-thirds of KSI cycle casualties occurred at or near a junction. They further reported (2009b) on the attribution of the contributory factors to the collision showing that, when considering all fatal and serious road collisions, the attributions to the cyclist or to the driver (non-cyclist) were fairly even. Only a relatively small proportion of contributory factors were attributed to both the cyclist and the driver. However, for young cyclists up to 24 years old, the proportion of contributory factors attributed to the cyclist (between approximately 55 and 80 per cent) was considerably higher than to the driver (between approximately 10 and 35 per cent).

The most common manoeuvres cyclists made prior to a collision were 'going ahead' and 'turning right'. While the cyclist was 'going ahead', a large proportion of cyclists KSI resulted from a vehicle turning right (15%), turning left (8%) or moving off/slowing down (6%). Frequent collision types involving a car and a bicycle were found to be due to a ‘car turning out of and into side road’, a ‘vehicle failing to stop at a junction’ and a ‘cyclist failing to stop at a junction’. Collisions due to a ‘cyclist crossing or entering road into path of vehicle’ were reported as being particularly frequent for child cyclists.

An analysis of contributory factors found that ‘failed to look properly’ was reported more frequently at junctions than away from junctions. The most frequent collision configurations that involved a car and a bicycle were a ‘car pulling out of side road’, a ‘cyclist crossing or entering road into path of vehicle’, a ‘vehicle failing to stop at a junction’, a ‘cyclist failed to stop at a junction’ and a ‘vehicle turning into side road’ (Knowles *et al*., 2009b, p. 20).

In a study by Johnson *et al.* (2011) travel direction, specifically turning left, was the greatest predictor of infringement by cyclists especially at junctions controlled by traffic lights: “Cyclists may perceive turning left to be a relatively safe manoeuvre since they are exposed to fewer points of conflict from cross traffic and cross traffic did have the deterrent effect and the perception of safety and opportunity to infringe decreased as the cross traffic volume increased” (p. 237).

A.5 Factors contributing to conflicts involving cyclists, particularly children

Knowles *et al.* (2009a) reveal that in over three-quarters of collisions in which a child cyclist was seriously injured, the child’s behaviour was reported as the primary contributory factor for the collision. It is not clear whether this means children are more likely than adults to behave in ways that result in a collision or whether the police are simply more likely to attribute contributory factors to a child. However, the crash circumstances of approximately 2,000 injured cyclists were examined and, for children under the age of 12 years, a high proportion was judged to be due to cyclist error. Furthermore, the 8 to 12-year-old cyclists were judged twice as likely to have caused a crash if they had no formal training. The two main contributory factors assigned to child cyclists involved in collisions were that the child ‘failed to look properly’ and ‘entered the road from the pavement’. (p. 35).

Miller (2012) reports that late detection of other road users leading to collisions has been highlighted as the most ‘basic driver error’. He cites one study in which ‘Looked-but-failed-to-see’ is common in vehicle–bicycle collisions and has been suggested as the likely cause of more than 50 per cent of crashes. Miller goes on to identify that drivers often report not being aware of cyclists they collide with, citing an in-depth study of bicycle–car accidents which reported that only 51 per cent of car drivers had noticed the cyclist prior to the collision. Detection of cyclists seems particularly poor when motor vehicles are pulling alongside the cyclist or approaching them from behind. A study of coroner’s records for fatal cycling accidents in London found that in collisions resulting from a motorist overtaking a cyclist, 44 per cent of drivers were unaware of the presence of the cyclist prior to the collision. This was the commonest crash configuration leading to fatalities.

A.6 Does training reduce the risk of injury for children?

To date, most evaluations of cycle training either focus on cycle training in the UK before Bikeability was introduced or on cycle training delivered in other countries.

In a study based on 818 questionnaires completed by Oxford hospital patients reporting the causes of accidents involving bicycles over a 12 month period, Simpson and Mineiro (1992) concluded that in the 8-12 years age group two-thirds of accidents were due to cyclist error. However, those who had had no formal cycle training were twice as likely to have caused the accident as trained cyclists. The authors recommended children in this age group 'should only be allowed on the roads after formal training' and that such training should become part of the National Curriculum.

In 1996, the Transport Research Laboratory (Savill *et al*., 1996) assessed whether 'cycle training schemes lead to improved, safer cycling skills and knowledge' for a group of 1,974 children. They were assessed at age 12; approximately two years after half this number had completed one of eight different training courses. Trained children performed significantly better than untrained children in the practical and knowledge tests. The most effective training courses were those with an on-road element and which were conducted over several weeks. Courses containing more than one stage, with each stage completed at different ages, were found to be effective too.

In September 2001, the Royal Society for the Prevention of Accidents published *The Effectiveness of Cyclist Training*, a review of 14 evaluations mostly undertaken by local authorities between 1976 and 1998. The results were mixed, but the review suggested the following:

- practical training impacts more on children's cycling than theoretical education
- on-road training is more effective than off-road proficiency testing
- learning in off-road or simulated training environments does not transfer to real on-road cycling
- the effect of training on children's cycling diminishes over time.

Other studies cast doubt on the efficacy of cycle training in improving the safety of children riding on the road in particular. One study, published in 2002, based on a survey of 336 children in two London schools, suggested that gender may be more important than training in explaining differences in cycling accidents, attitudes and behaviour (with girls more likely to adopt 'safe attitudes' and boys more likely to 'show off') (Colwell and Culverwell, 2002).

An evaluation of the Australian 'Bike Ed' school cycle training programme (Carlin *et al*., 1998), based on interviews with 148 cases of children in hospital emergency departments with cycling injuries and 130 controls, drew the following conclusion:

> *This educational intervention does not reduce the risk of bicycle injury in children and may possibly produce harmful effects in some children, perhaps*

due to inadvertent encouragement of risk taking or of bicycling with inadequate supervision.

(p. 22)

(It should be noted that Bike Ed recommends children under the age of 12 should not ride on the road, and children over the age of 12 only with adult supervision.)

Yet in 2010, Ipsos MORI reported that children overwhelmingly felt their abilities to judge risk improved following Bikeability training, a finding confirmed by parents. Parents felt children's safety on the road had improved and children felt more confident cycling on the road.

Moreover, a recent study from Belgium (Ducheyne *et al*., 2013) reported short-term effects of practical cycle training on basic bicycle handling skills, with training delivered in a traffic-free environment for children in five primary schools (two controls). Participating schools were assigned to an intervention or control group and cycling skills were assessed using a practical assessment at baseline and immediately after the intervention. The authors found that "gender, social and economic status and initial cycling skills level had no significant influence on the effects of the cycle training" (p. 38). However, the cycle training was found to have a statistically significant effect, with an effect size of 1.30, on children's cycling skills.

In 1994, van Schagen and Brookhuis published the results of an investigation into children's performance of 'motor task' and 'cognitive task' components associated with cycle training. The first roughly equates to Bikeability level 1 skills (e.g. balancing, controlling the bicycle, making it go where you want it to, etc.), the second consist of retrieval and application of (a) 'behavioural rules' (e.g. signalling, road position, etc.), and (b) 'rules that regulate traffic interactions in dynamic situations' (e.g. priority situations), and together roughly equate to Bikeability Level 2 skills.

Both approaches had an equally positive effect on simple behavioural strategies, such as signalling and visual search behaviour. Correct application of priority rules appeared to be very difficult to teach. There was no effect of either of the two training methods. It seemed that children apply informal rules rather than formal rules when dealing with other traffic. It is hypothesized that these informal rules should form the starting point for training activities, because formal rules do not fit into children's cognitive framework of schemes, and therefore cannot be stored and retrieved effectively.

(p.223)

McLaughlin and Glang (2010) investigated the "Bike Smart" programme, an eHealth programme that teaches bicycle safety behaviours to young children in the United States. Two groups, intervention and control, took assessments which included computer-based knowledge items (e.g. safety rules, hazard discrimination) and a behavioural measurement of helmet placement. The results revealed that regardless of gender, cohort or grade, the participants in the treatment group exposed to the Bike Smart programme showed greater gains that those in the control group. The effect sizes were 2.05 on hazard discrimination and 1.42 on safety rules.

Taken together, previous studies reveal a diversity of different approaches to cycle training (e.g. theoretical education, practical training in traffic-free, simulated and on-road environments) and evaluation methods and data sources (e.g. knowledge tests, observed demonstrations, household surveys, hospital patient questionnaires). Where effect sizes are reported, these tend to be high in educational terms although, for the studies cited, these are only measured over the short term.

Appendix B – On-screen quiz

B.1 Question development

Two sets of questions were developed – one to find out background information about the participants and their cycling behaviour and one designed to assess their hazard perception and appropriate response ability.

The first set of questions sought to find out information such as:

- gender
- mode of travel to school
- length of journey to school
- mode of travel to other places e.g. the park, going shopping etc
- family access to a car
- bicycle ownership
- cycling enjoyment
- cycling experience (e.g. on roads, off road/pavements)
- frequency of cycling
- cycling confidence
- who they cycle with.

The questions for inclusion in the on-screen quiz were developed to address the four areas of observation, signalling knowledge and skills, knowledge of priorities and road position and were pitched to match the Level 2 National Standard criteria.

We aimed to include a variety of item types to ensure the quiz would be varied and interesting for participants and to make the most of the multi-media approach. For this reason, lots of photographs and video clips were included to show realistic, credible situations that children may find themselves in when cycling.

The majority of questions were multiple-choice, requiring pupils to click on one or more answers. In most cases, pupils were also given the option to answer 'I don't know' (in a paper test, pupils tend to leave questions blank that they do not want to or cannot answer and this was not an option in the on-screen version).

B.2 The on-screen quiz

Following question writing, review and refinement, they were inputted into Questback, the online delivery platform, in order to create an on-screen version of the quiz. This was informally trialled with a small number of 9-11 year olds to check that it functioned appropriately and that the questions were pitched suitably in terms of readability / understanding.

Figure B1-B6 show a number of screen shots of the questions that appeared in the quiz. The full set of questions is provided in Appendix E.1.

Figure B1 Example screen shot – introduction to the quiz

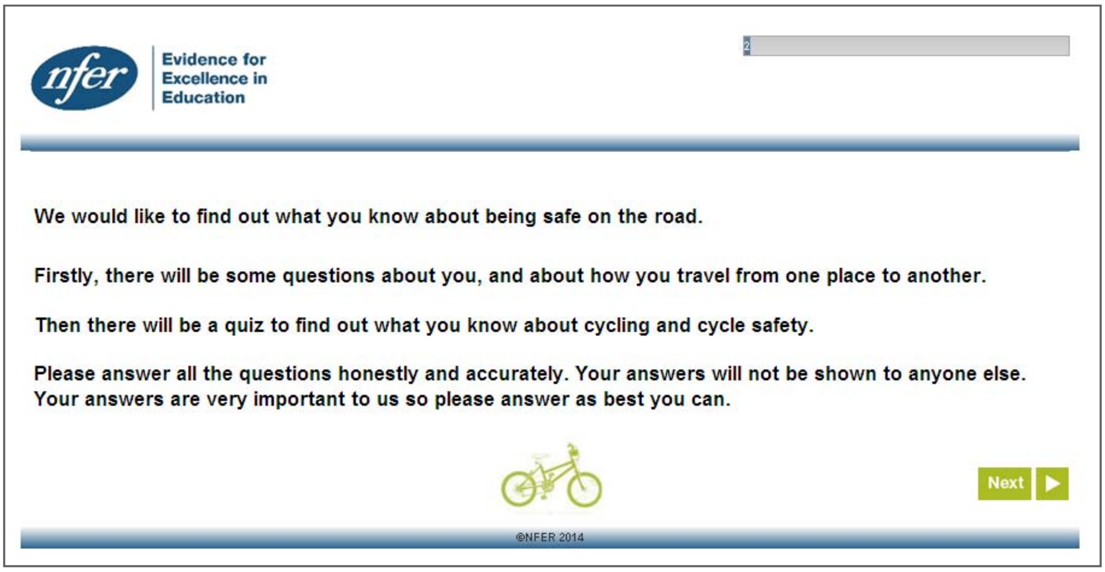


Figure B2 Example screen shot – background data collection

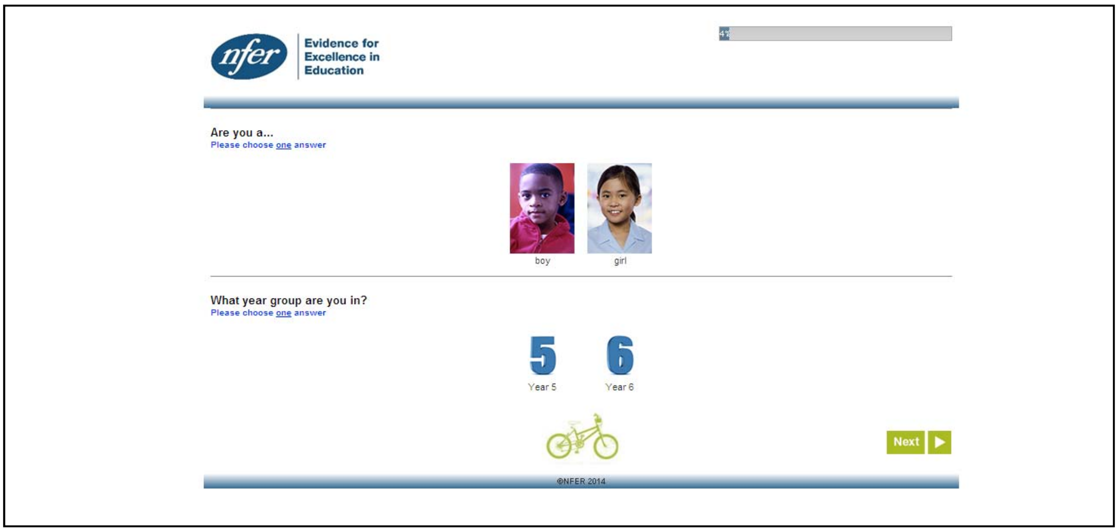


Figure B3 Example screen shot – multiple choice

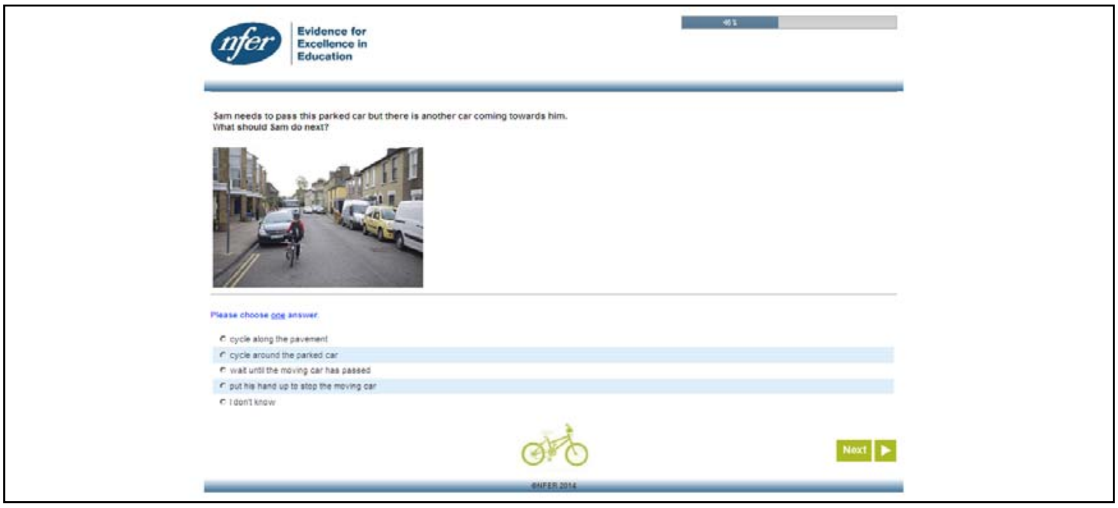


Figure B4 Example screen shot – film clip stimulus

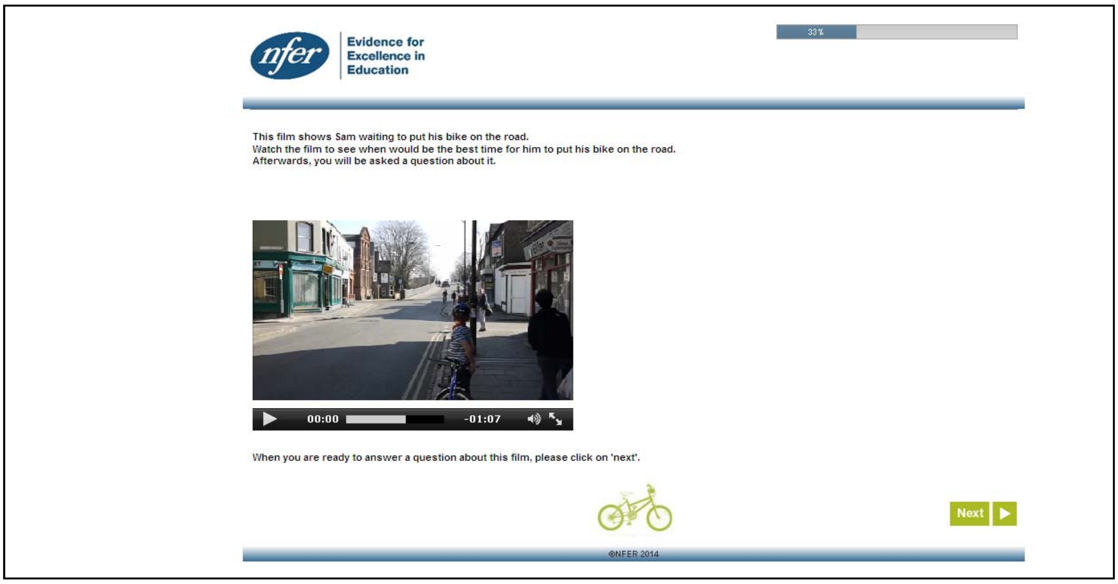


Figure B5 Example screen shot – sequencing

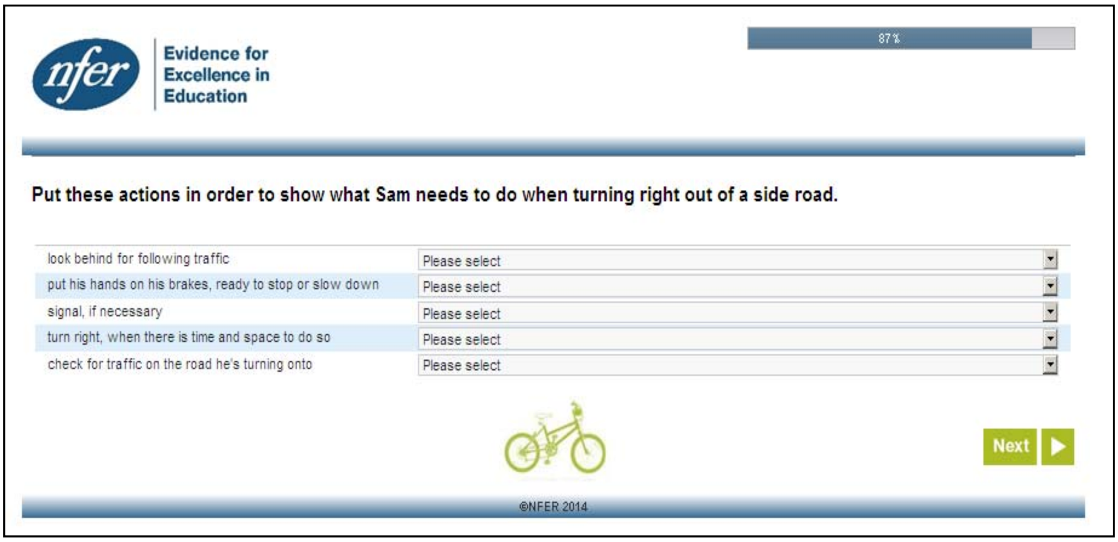


Figure B6 Example screen shot – table completion

When schools agreed to participate in the study, the Year 5 class teacher was requested to complete a Pupil Data Form identifying the name, gender and date of birth of pupils participating in the on-screen quiz. In the intervention schools, teachers were also asked to identify those pupils due to partake in the Bikeability training in the summer term. Each school was then sent an individual log-in for each pupil and they were assigned to either Quiz 1 or Quiz 2. Within a class, equal numbers of pupils were assigned to Quiz 1 and Quiz 2; individuals took alternating quizzes at each time point.

Appendix C – Practical assessment

C.1 Assessment development

The purpose of the practical assessment was to gather data about trained pupils in a practical scenario which could then be mapped to their data from the on-screen quiz to see if there was a correlation between the two scores and thus provide validation of the on-screen test (i.e. to identify whether or not it can be used as a predictor of pupils' cycling proficiency/hazard awareness).

The practical assessment was designed by a National Standard Instructor Trainer (NSIT) who is very familiar with the requirements of the Level 2 National Standards.

Rather than replicate the assessment of all Bikeability Level 2 outcomes, the assessment sought to gather information about four domains aligned with the on-screen quiz: observation, communication, road position and priorities. Pupils were required to demonstrate their competence, confidence and consistency in each of these four domains whilst completing two drills twice:

Drill 1: Children followed a 'lozenge' circuit on the major road[1], passing the minor side road, a parked car, and performing U-turns at the ends of the circuit.

Drill 2: Children followed a 'bent sausage' circuit starting on the major road, turning left into the minor side road, performing a U-turn before turning right into the major road and ending with a U-turn.

In order to ensure consistent and reliable assessments, an easy to apply four-point rating scale was devised for recording observations of the four domains of interest:

0 = not seen by the assessor

1 = rarely seen by the assessor

2 = mostly seen by the assessor

3 = always seen by the assessor.

The assessment was delivered by experienced NSIQs who routinely deliver Level 2 Bikeability training. They assessed children that they had not trained themselves.

The practical assessment was piloted and refined before being used with all the participating schools. Training was provided to the NSIQs before they took on their assessment role. As far as possible, assessments were arranged within two weeks of Bikeability training being completed in each school in the summer term (mostly in July 2014), and then within a four week period in the first half of the autumn term (mostly in October 2014).

C.2 Preparing for the assessment

C.2.1 Risk controls

Risks to children participating in the practical assessments were minimised with the following controls:

---

[1] In the National Standard, a 'major road' is any road with continuing traffic and where a 'minor road' terminates (for example, a T-junction). A major road is not necessarily busier than a minor road.

- only including in the assessment children who had passed Bikeability Level 2 training
- obtaining informed and active parental consent on the basis of children's prior cycling experience
- ensuring that each child was aware that they could tell the assessor that they would like to miss out elements of the assessment that they were worried about
- only including in the sample children who had roadworthy bikes and the assessors were confident had the bike handling skills required for cycling on the road before starting the practical assessment
- excluding the most challenging Bikeability Level 2 drill (right hand turn in from a major to a minor road) from the on-road assessment
- risk assessing the on-road assessment sites in accordance with standard Bikeability site risk assessment requirements.

C.2.2 Site selection

The on-road assessment took place with a group of up to 12 trainees accompanied by two instructors (in line with the 1:6 ratio allowed for Bikeability Level 2 delivery). A typical first day Bikeability Level 2 T-junction was selected near the school with good sight lines, some parked cars and light traffic, and ideally road markings. Assessors completed the standard Bikeability Level 2 site risk assessment, noting static and dynamic hazards observed during the assessment.

Assessors assessed the site they used and a back up site near the school before the assessment session began, and completed a risk assessment sheet for both sites (please refer to section C.4). The risk assessment sheet included a section on common potential hazards for children that the assessors rated as low, medium or high, and space was provided for comments on any changes to these ratings during the session (e.g. traffic becoming heavier or lighter as the session progressed). They noted the position of the pupils and the assessors at the site.

Photographs of typical sites for carrying out the assessments are shown in Figures C1 - C4 below.

Figure C1 T-junction

Figure C2 Major road / Minor road

Figure C3 Passing a minor road and cars

Figure C4 Passing a minor road

C.2.3 Pupil preparation

At the assessment site, pupils were assigned to two groups of up to six pupils, with each group supervised by one assessor. They were walked or 'snaked' to the drill site in the same manner as for Bikeability training.

C.3 Carrying out the assessment

All children attempted Drill 1 and Drill 2 twice. When all children in both groups had completed their first drill twice, the groups swapped and they were assessed by the other assessor while they did the other drill twice. Children in each group took turns to complete each drill without instruction and alone, unaccompanied by any other children from their group. Waiting children could observe but were not allowed comment on or discuss their peers' performance.

Each assessor was responsible for supervising and assessing up to six children performing one drill. Assessors positioned themselves where they could see all aspects of each child's performance in order to assess their skills in observation, communication, road position and priorities.

Both groups performed the drills in parallel. Both instructors observed and assessed each child performing their drill, and recorded their assessment after each child had completed their two turns on the score sheets provided (see section C.5).

Each practical assessment took up to 120 minutes, for 12 children, in total.

C.4 Site risk assessment form

Evidence for Excellence in Education

Bikeability Hazard Perception Practical assessment

Site risk assessment

School:		Assessment Date:	
School First Aider:		School Telephone:	
First assessor:		Second assessor:	

Date & time of site risk assessment	
Major road	
Minor road	
Sketch site map with position of assessors and pupils Attach site photographs	
Sketch route map from school to site	

Potential hazards	Lower risk	Higher risk	Rating	Comments
Road width	Wide	Narrow		
Road bends	Straight	Bending		
Sight lines	Uninterrupted	Interrupted		
Road surface	Smooth	Rough		
Parked vehicles	Few	Many		
Traffic level	Light	Heavy		
Major rd markings	Present	Absent		
Minor rd markings	Present	Absent		
Weather	Mild and dry	Hot/inclement		
Other				

Notes:
(1) risk to children rating scale – 1 = low risk, 2 = medium risk, 3 = high risk
(2) assessors to comment on any changes to these ratings during the session.

C.5 Practical assessment form

Practical assessment score sheet

Assessor's name ………………………………………….

School name ………………………………………….....

Date of assessment ……………………………………..

observation – are they aware of what is around them at all times?

communication – do they know how and when to communicate?

road position – do they take the best road positions for their journey?

priorities – do they know priorities and how to assert them?

Scoring: 0=never seen 1=rarely seen 2=mostly seen 3=always seen

GROUP A	OBSERVATION	COMMUNICATION	ROAD POSITION	PRIORITIES
Name of child	0-3	0-3	0-3	0-3
1				
2				
3				
4				
5				
6				

GROUP B	OBSERVATION	COMMUNICATION	ROAD POSITION	PRIORITIES
Name of child	0-3	0-3	0-3	0-3
1				
2				
3				
4				
5				
6				

Appendix D – Recruitment

D.1 Age group and pupil selection

In order to assess the immediate and longer-term impact of Bikeability training on children's hazard perception and appropriate response ability, it is important to test children both at the point of training and also some months later. Due to the desirability of carrying out the practical assessment in the warmer months, the ideal period for carrying out the research was during the summer term, repeating the assessment a minimum of two months later, in September 2014.

The project involved pupils who were in Year 5 (Y5) in the summer term and who moved into Year 6 (Y6) in September 2014. The rationale for choosing Year 5 pupils was as follows:

- there is no evidence of any systematic differences in the ability of Y5 versus Y6 children in perceptual abilities or in hazard detection. Across any two adjacent school years there is inevitably a mix of ages, so children in any two adjacent years may be only a few months apart in terms of age.
- testing Y6 pupils in the summer term with follow-up testing in September (during Y7) would involve gaining the permission of pupils' secondary schools as well as administrative issues (tracking the destinations of pupils) and logistical problems in arranging the Y7 practical assessments at a central location. In addition, cycling behaviour may alter as pupils move on to secondary schools. Although this would be the same for both trained and untrained pupils, any such difference may be conflated with the specific effects of Bikeability, making the longer term impact of training more difficult to determine.
- approximately one-third of Bikeability training is carried out with Y5 pupils, therefore there is sufficient Y5 training during the summer term to recruit sufficient schools / pupils to accomplish the research effectively. The comparison group of pupils will be untrained Y5 pupils from schools in the same areas where training is delivered in Y6.

The research design proposed:

- a target sample for the on-screen test of approximately 1000 Year 5 pupils
- approximately half of these (500 pupils) would be from 20 schools undertaking Y5 Bikeability training in the summer term
- from the pupils participating in the training, a sub-sample (200-240 pupils) would be randomly selected to undertake the practical assessment in addition to the on-screen quiz

- the other half of the on-screen sample will be Y5 pupils from 20 schools that carry out Bikeability training in Year 6
- the on-screen sample would be likely to include children with varying degrees of competence and experience in cycling.

D.2 Approach to schools

Training and comparison schools were approached through the local Bikeability schemes who recruit them for training purposes. Bikeability schemes have lists of schools in their areas which identify the number of pupils in each school year to be trained during the summer.

Nine Bikeability schemes (four local authority areas [Areas 1-4] plus five London regions [Area 5]) provided the names of the schools that had training planned with Y5 and Y6 pupils. These schools were approached to seek their interest in taking part in the research – either as a 'trained' or 'comparison' school. 'Trained' school groups included Y5 pupils who received Bikeability Level 2 training during the summer term. 'Comparison' school groups included Y5 pupils who did not receive Bikeability training (because the training was provided to their Y6 peers).

Schools were successfully recruited from six schemes. In addition to these schools, a top up sample was also drawn to gather further data from the on-screen quiz. A random selection of schools from across England were invited to participate in a one-off assessment in order to provide information about the functioning of the on-screen quiz.

A summary of the numbers of schools recruited and completing the assessments at each of the three time points are shown in Table 2.1 in Section 2.5 of the report.

Appendix E – Analysis

E.1 On-screen quiz analysis

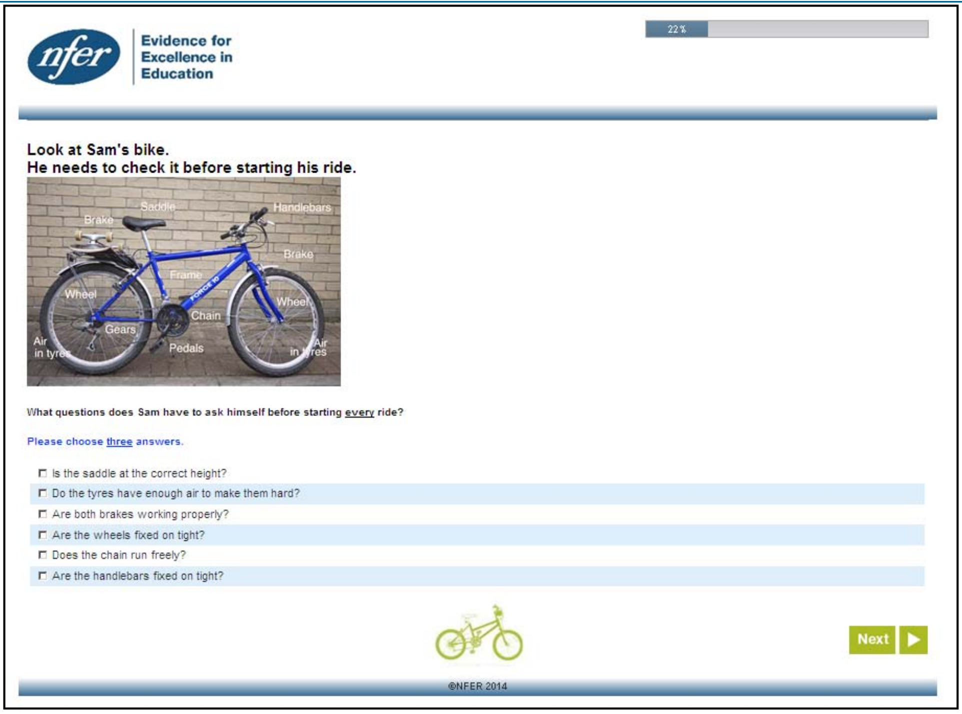
This analysis included item level statistics (relating to each individual question) and analysis of gender differences for each test. Outcomes from this analysis have been used to inform the selection of items for final analysis i.e. by informing decisions about which items did not function as anticipated and removing these from the final analysis.

This section will focus upon the functioning of individual quiz items. For each question in the quiz, a table displays the following information:

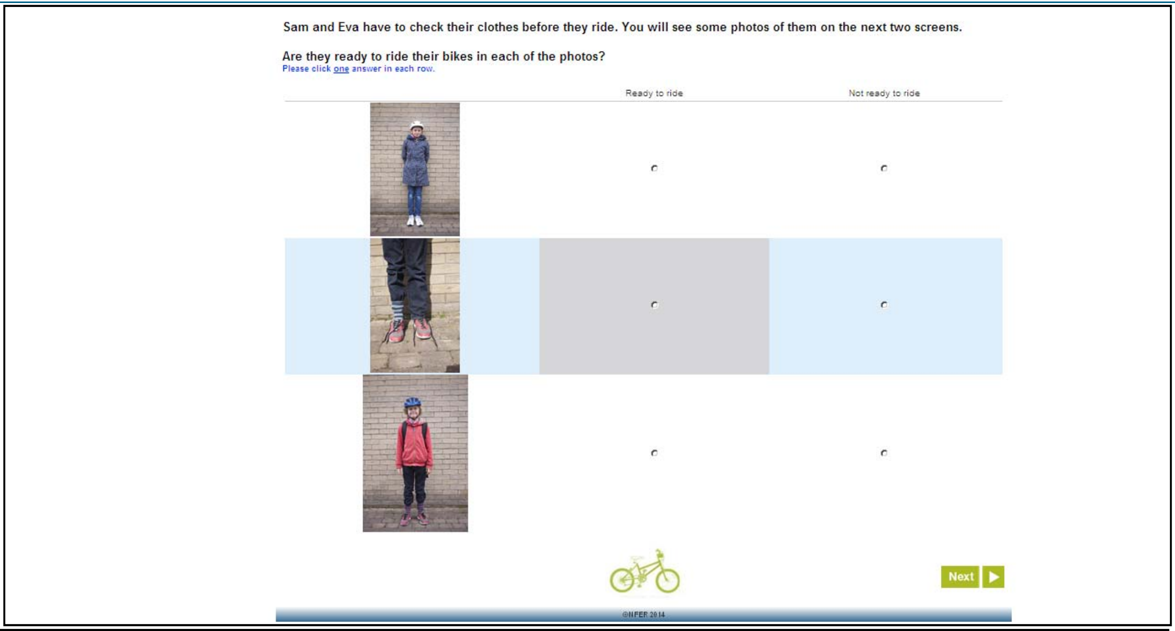
- maximum number of marks available – either 1 or 2 marks depending on the complexity of the question
- mean score
- overall facility - the mean mark divided by the number of marks available for the item
- discrimination index - the extent to which the item measures the same construct as the rest of the test and therefore how well it discriminates
- IRT slope - the slope of the curve helps to determine the usefulness of the item. In general for tests of literacy or numeracy, less than 0.4 is considered 'uninformative', i.e. the item does not tell us much about the pupils' ability. However, we might expect lower slopes in a test of hazard perception and appropriateness of response because the underlying level of knowledge (which is what we are discriminating) is low to begin with.
- domain – questions have been allocated to observation, communication, road position or priorities (or Level 1)
- common – a total of 15 questions appeared in both versions of the quiz allowing us to link the test data
- notes – these include details of the DIF analysis* outcomes and any other pertinent information about the functioning of the question.

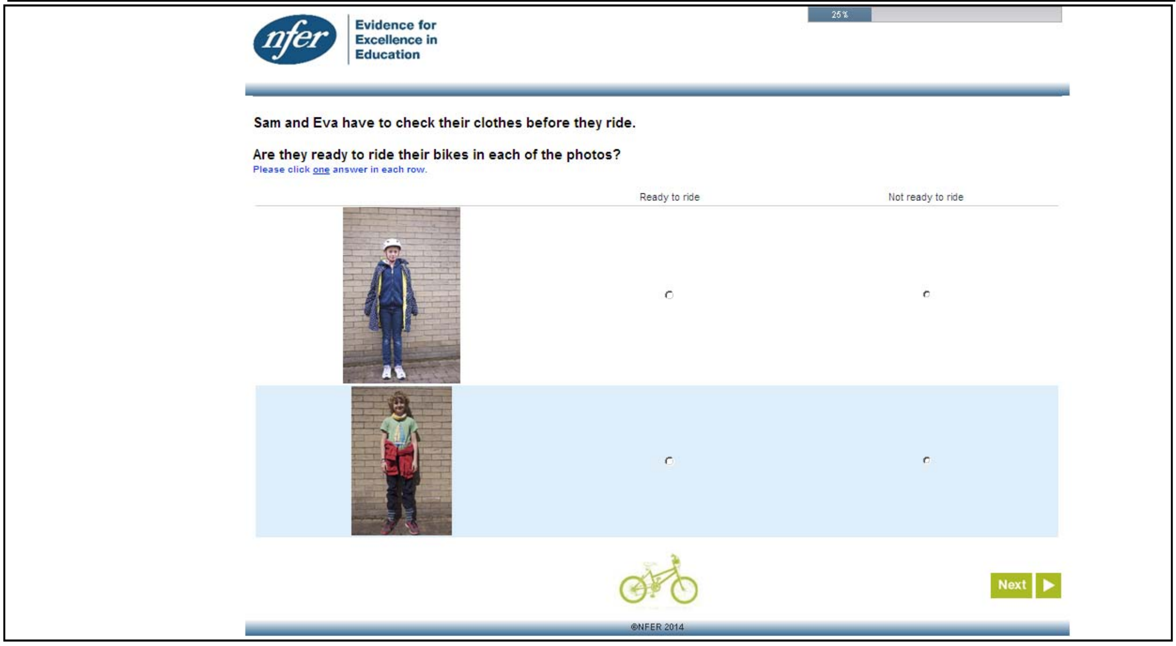
* For each test an analysis of differential item functioning (DIF analysis) was performed to ascertain whether any particular items seemed to favour boys or girls, taking into account the overall difference in their performance. This analysis identifies items where the pattern of performance is significantly different from the overall pattern for the test. Items identified in this analysis are not necessarily biased towards one group or another; the results may reflect genuine differences in performance. The analysis gives indications as to which items may need further investigation. The level of significance gives the probability (*p*) that this result occurred by chance and there is no real difference, so the smaller the figure, the higher the significance (i.e. $p<0.0001$ is the highest level of significance generally recognised) and the more confident we can be that the effect is genuine. Results of a significance test are indicated by the following:

Probability	Percentage level	Probability	Percentage level
$p<0.05$	5% level	$p<0.001$	0.1% level
$p<0.01$	1% level	$p<0.0005$	0.05% level

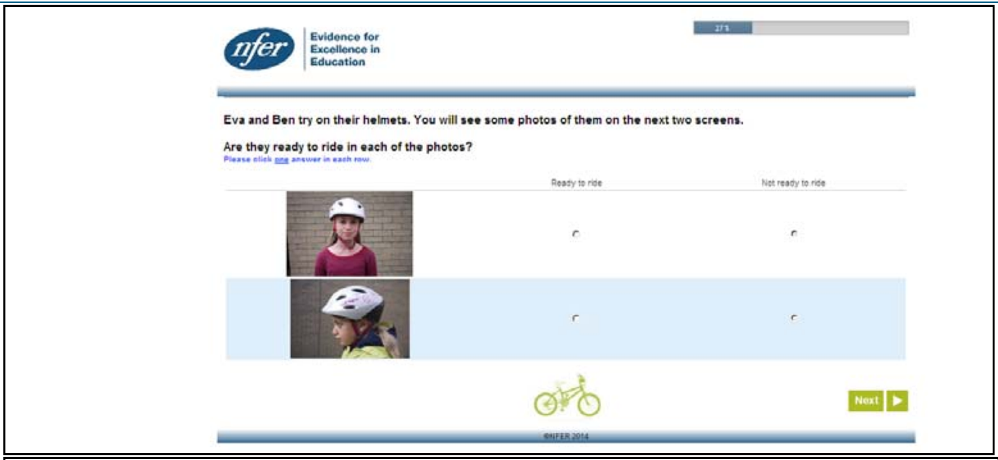


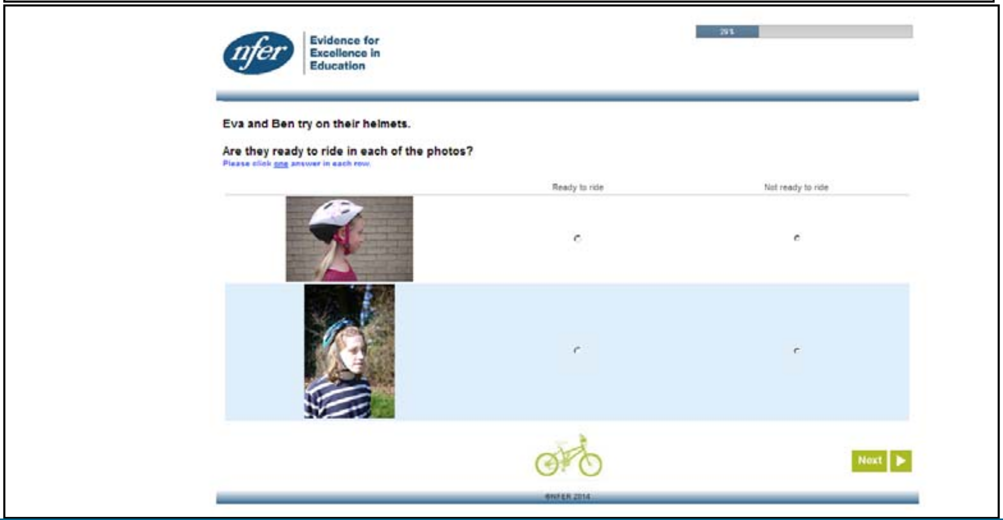
		Test 1 – Item 1					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	1.04						
Facility	52	52	40	76	52	68	55
Discrimination	0.34						
IRT Slope	0.53						
Domain	Level 1						
Common item	×						
Notes	Overall: 52% obtained 1 mark; 26% obtained 2 marks						



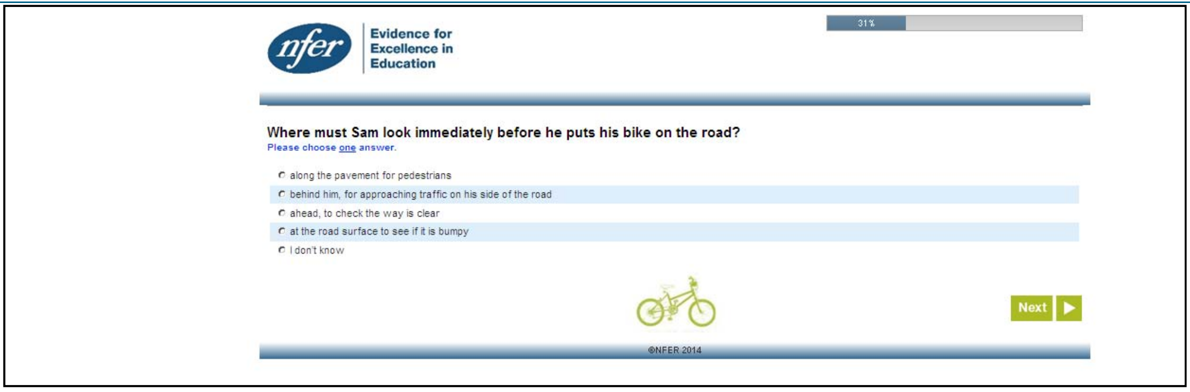


		Test 1 – Item 2					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.73						
Facility	73	78	70	79	80	67	75
Discrimination	0.09						
IRT Slope	-						
Domain	Level 1						
Common item	✗						
Notes	Item removed due to poor functioning.						

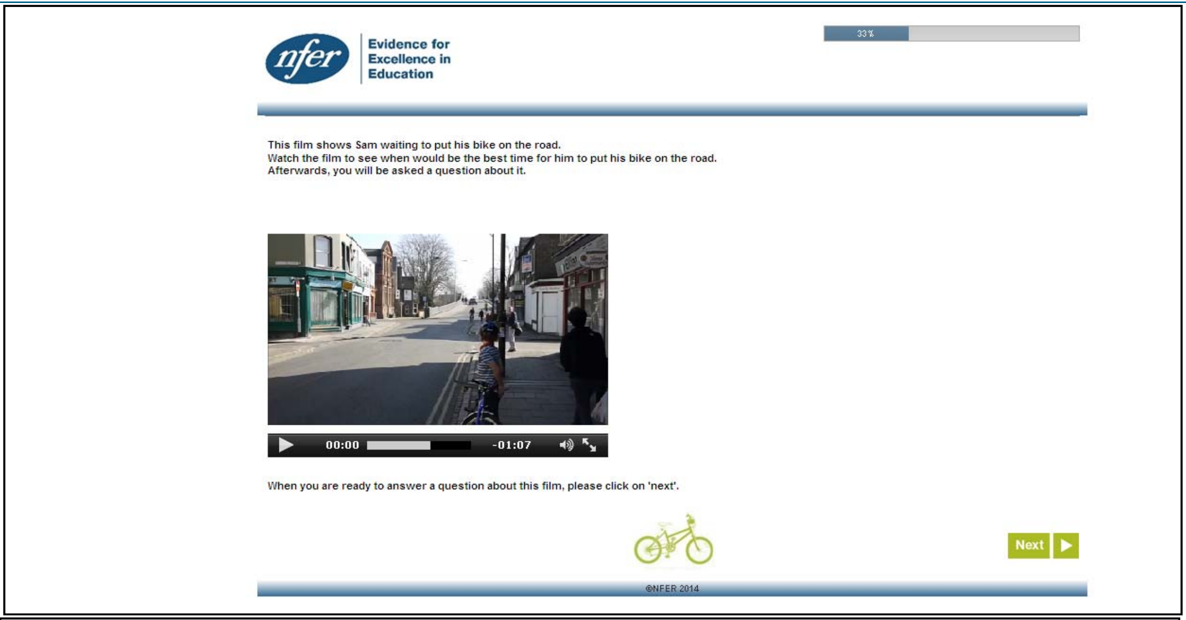


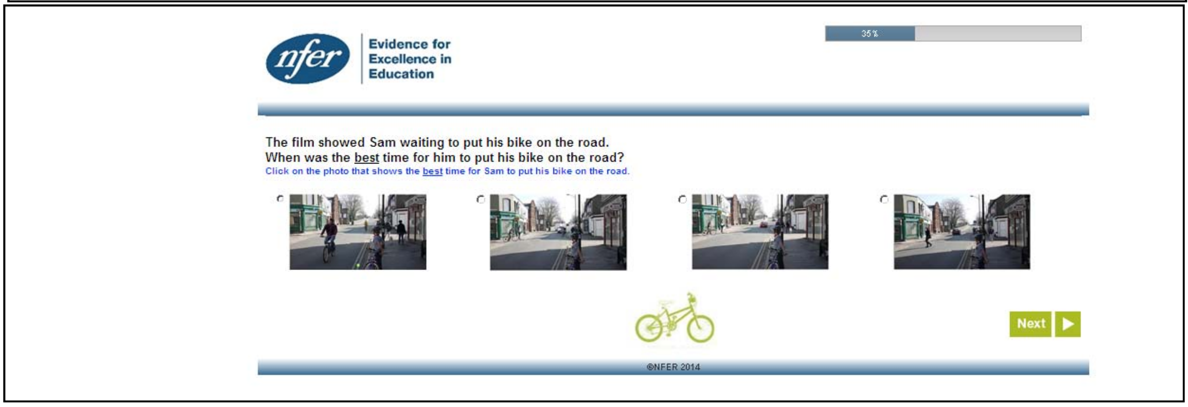


		Test 1/2 – Item 3					
		Phase 1 (Baseline)		**Phase 2 (Post-training - June)**		**Phase 3 (Post-training – September)**	
	Overall	**Trained**	**Comparison**	**Trained**	**Comparison**	**Trained**	**Comparison**
Max score	1						
Mean score Test 1 Test 2	0.73 0.75						
Facility Test 1 Test 2	73 75	69	70	82	85	86	71
Discrimination Test 1 Test 2	0.26 0.30						
IRT Slope	0.49						
Domain	Level 1						
Common item	✓						
Notes	Favours girls in both versions of the test (T1: 5% level of significance; T2: 0.1% level of significance). The third part of this question did not operate as anticipated. Analysis is based responses to the question, excluding this element (i.e. if used in future, the third photograph would need to be removed).						



		Test 1/2 – Item 4					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	0.66 0.68						
Facility Test 1 Test 2	66 68	58	59	90	69	86	67
Discrimination Test 1 Test 2	0.40 0.37						
IRT Slope	0.80						
Domain	Observation						
Common item	✓						
Notes							

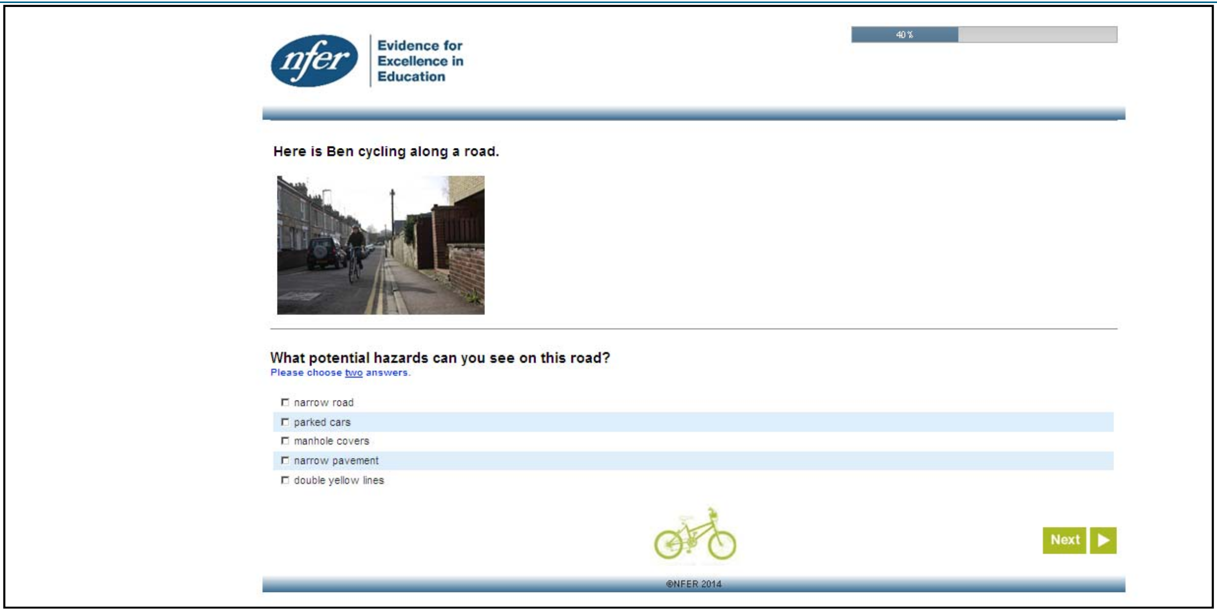




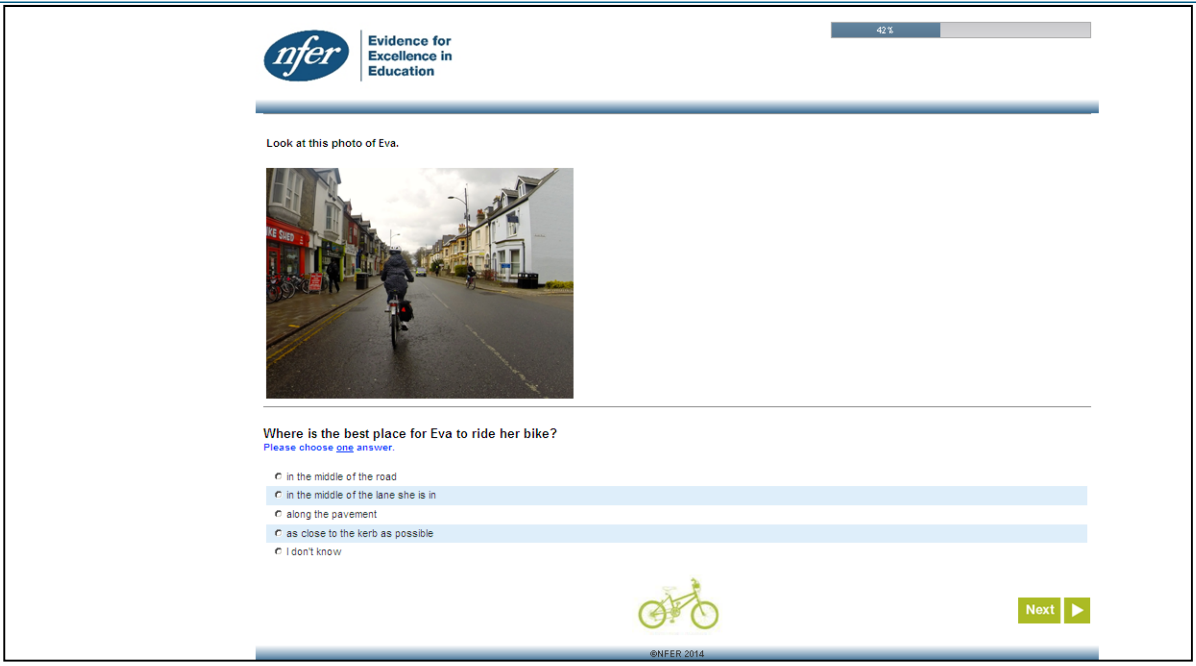
		Test 1 – Item 5					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.72						
Facility	72	71	66	92	62	83	76
Discrimination	0.33						
IRT Slope	0.66						
Domain	Observation						
Common item	✘						
Notes							

		Test 1 – Item 6					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.76						
Facility	76	71	70	95	73	83	79
Discrimination	0.34						
IRT Slope	0.69						
Domain	Observation						
Common item	×						
Notes							

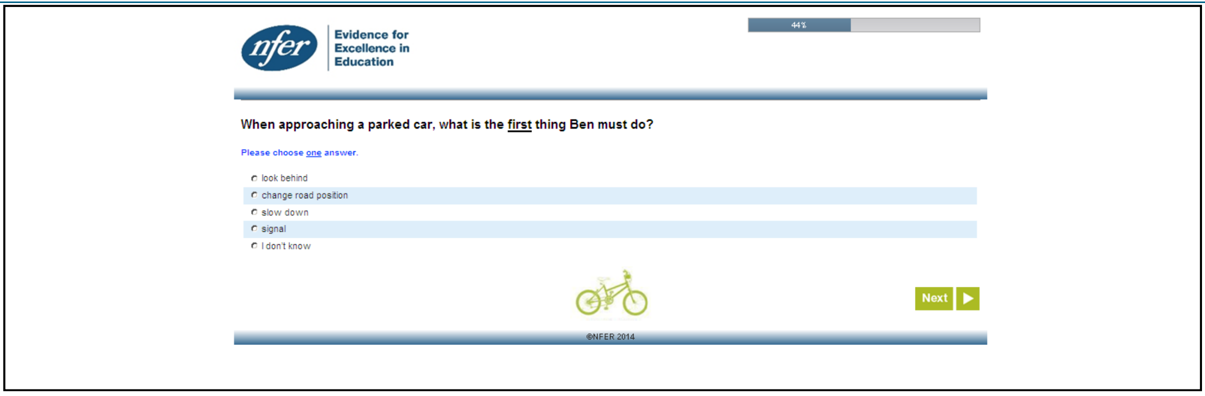
		Test 1 – Item 7					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.25						
Facility	25	16	16	45	27	41	27
Discrimination	0.27						
IRT Slope	0.49						
Domain	Road position						
Common item	×						
Notes							



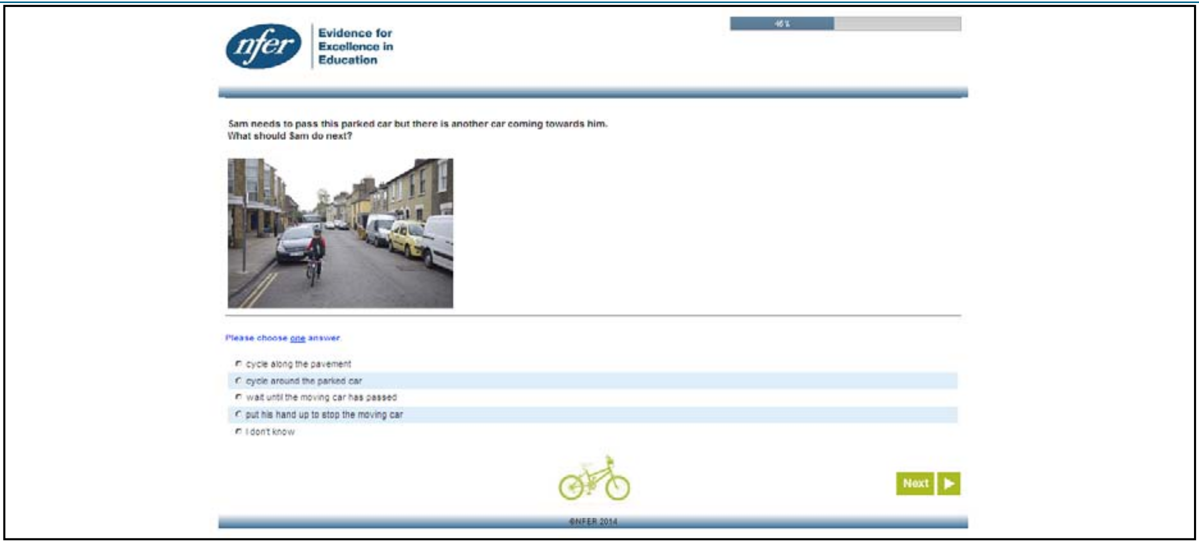
		Test 1 – Item 8					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.89						
Facility	45	40	38	67	46	56	43
Discrimination	0.25						
IRT Slope	0.34						
Domain	Observation						
Common item	×						
Notes	Overall: 21% obtained 1 mark; 34% obtained 2 marks						



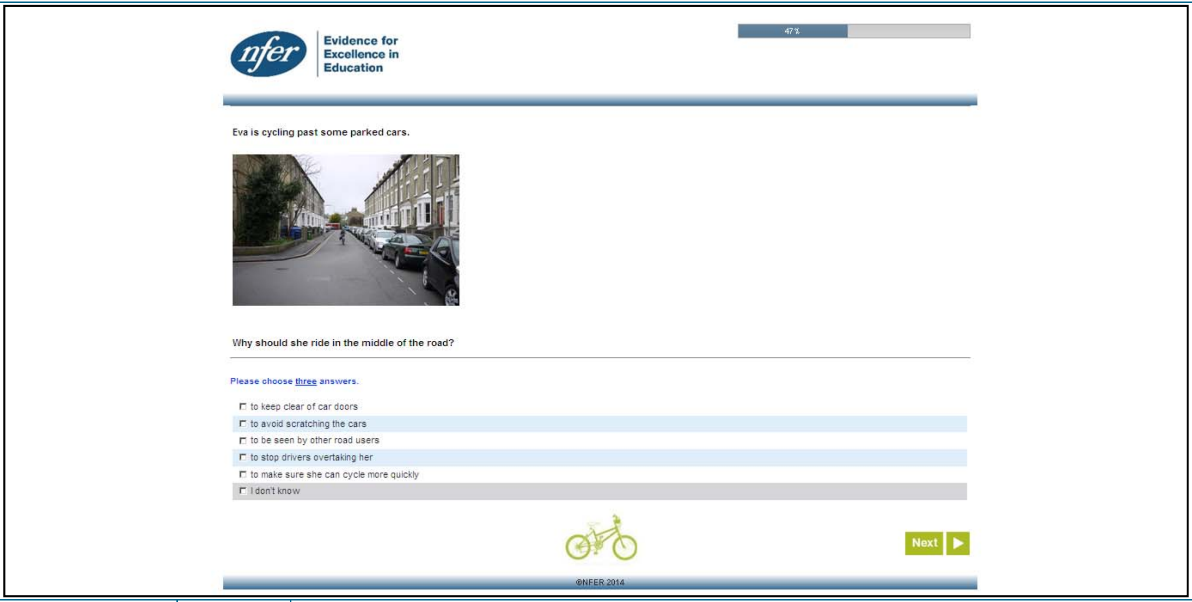
		Test 1/2 – Item 9					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	0.24 0.19						
Facility Test 1 Test 2	24 19	6	16	52	25	52	8
Discrimination Test 1 Test 2	0.30 0.38						
IRT Slope	0.66						
Domain	Road position						
Common item	✓						
Notes							



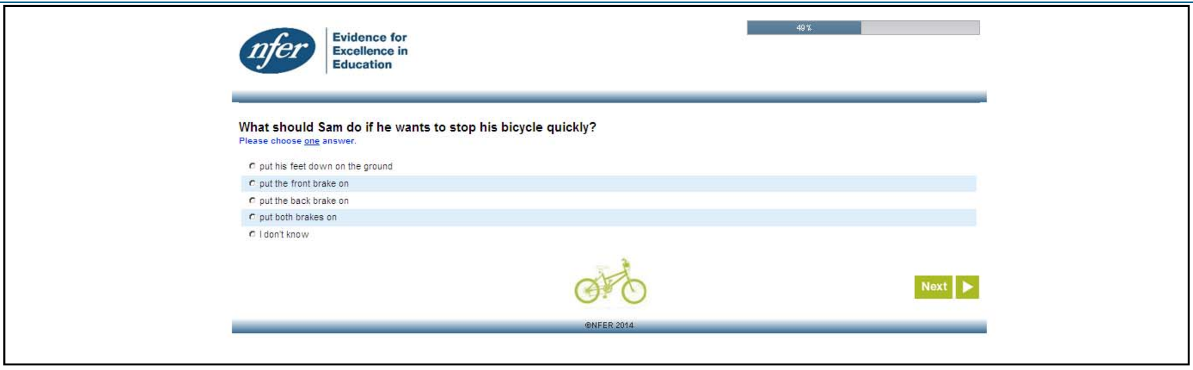
		Test 1/2 – Item 10					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	**Trained**	**Comparison**	**Trained**	**Comparison**	**Trained**	**Comparison**
Max score	1						
Mean score Test 1 Test 2	0.27 0.25						
Facility Test 1 Test 2	27 25	15	23	49	22	36	26
Discrimination Test 1 Test 2	0.17 0.16						
IRT Slope	-						
Domain	Observation						
Common item	✓						
Notes	Item removed due to poor functioning.						



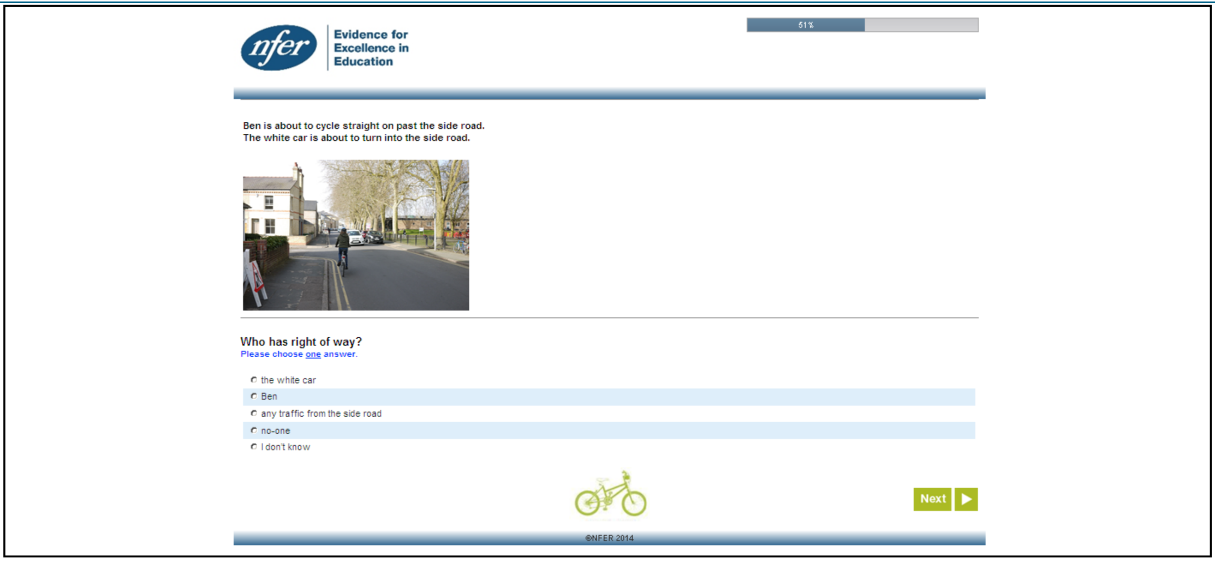
		Test 1 – Item 11					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.72						
Facility	72	65	61	90	69	93	80
Discrimination	0.34						
IRT Slope	0.66						
Domain	Rights of way						
Common item	✗						
Notes							



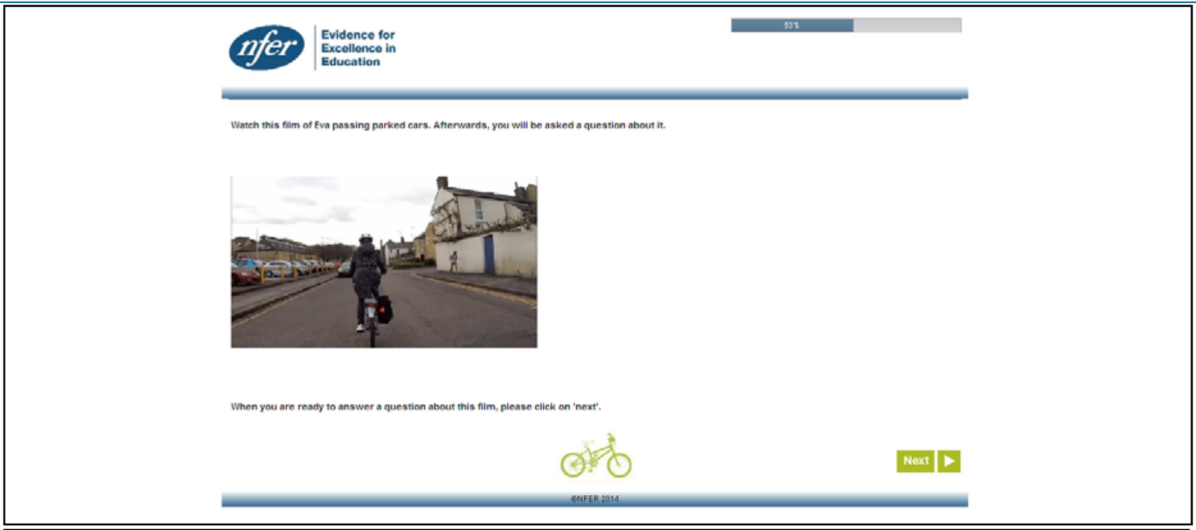
		Test 1 – Item 12					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.93						
Facility	47	43	39	72	46	62	42
Discrimination	0.45						
IRT Slope	0.81						
Domain	Road position						
Common item	✗						
Notes	Overall: 67% obtained 1 mark; 13% obtained 2 marks						

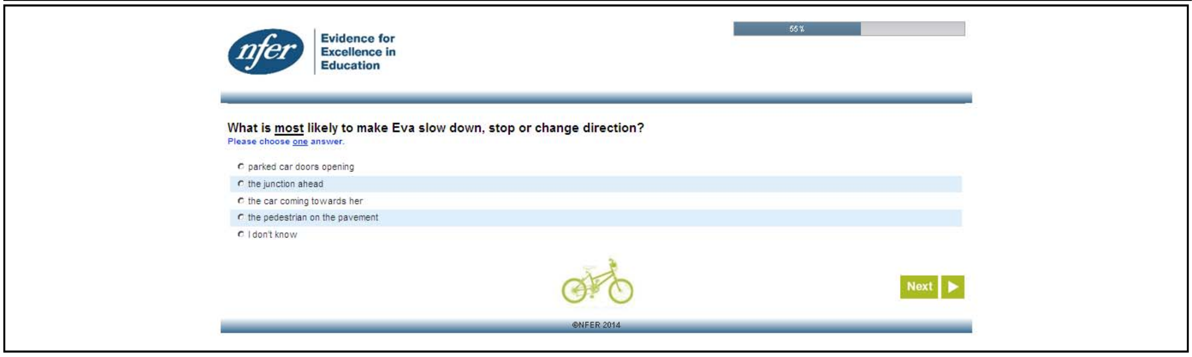


		Test 1/2 – Item 13					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	 0.68 0.68						
Facility Test 1 Test 2	 68 68	65	68	70	74	74	63
Discrimination Test 1 Test 2	 0.22 0.25						
IRT Slope	0.36						
Domain	Level 1						
Common item	✓						
Notes							

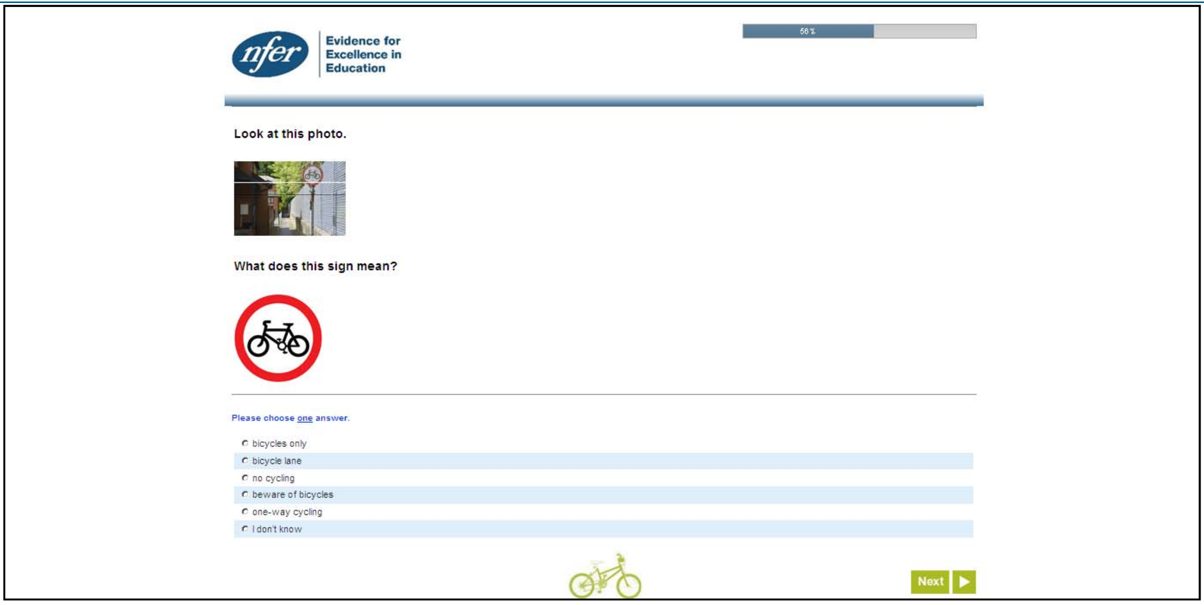


		Test 1 – Item 14					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.43						
Facility	43	38	39	66	33	58	36
Discrimination	0.21						
IRT Slope	0.31						
Domain	Rights of way						
Common item	✗						
Notes							

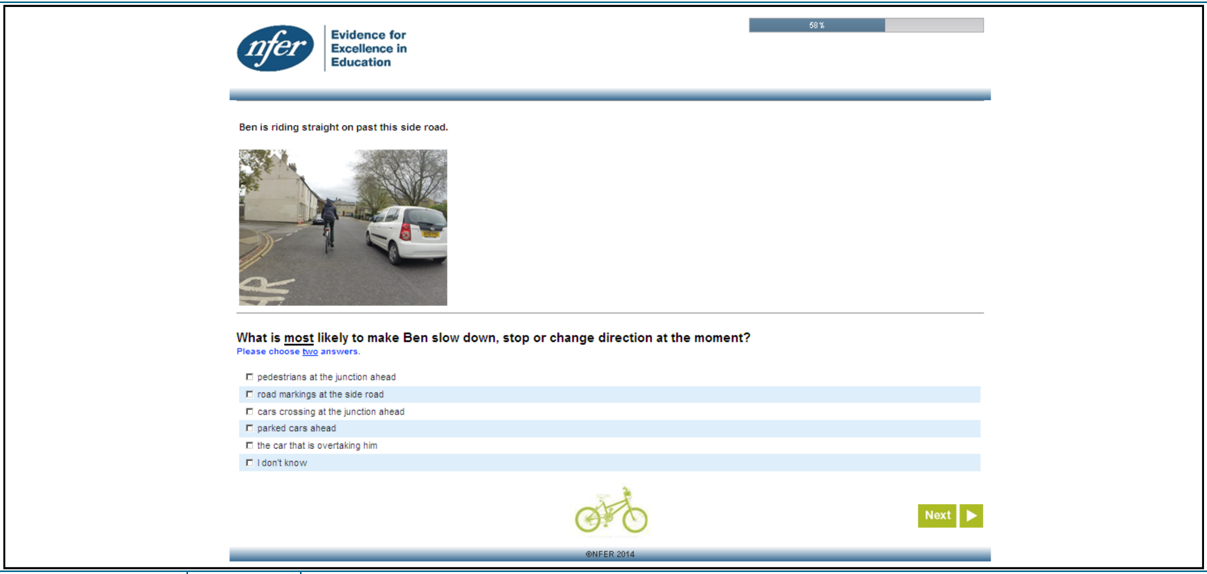




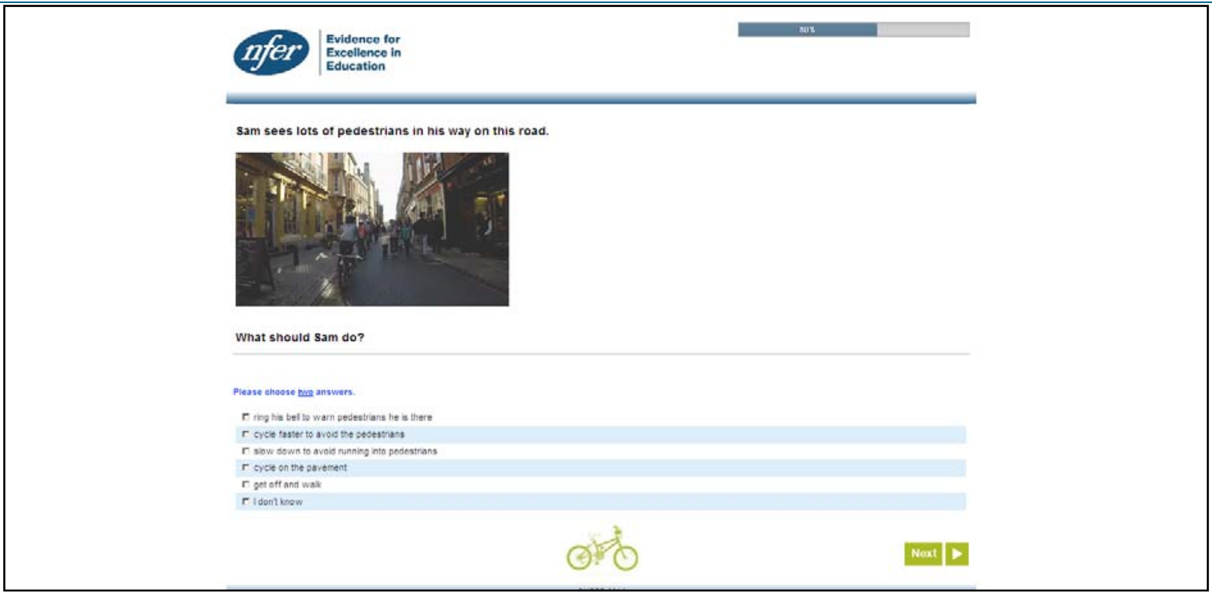
		Test 1 – Item 15					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.54						
Facility	54	53	57	42	44	49	60
Discrimination	0.04						
IRT Slope	-						
Domain	Observation						
Common item	✗						
Notes	Item removed due to poor functioning.						



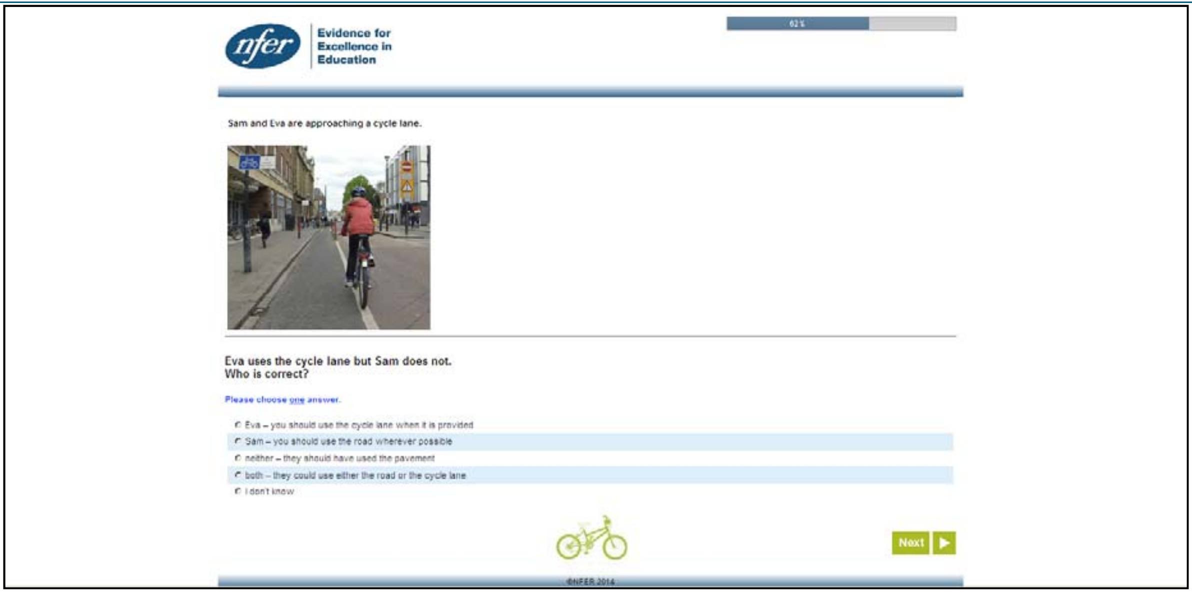
		Test 1 – Item 16					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.13						
Facility	13	7	9	45	11	20	6
Discrimination	0.22						
IRT Slope	0.49						
Domain	Rights of way						
Common item	✗						
Notes							



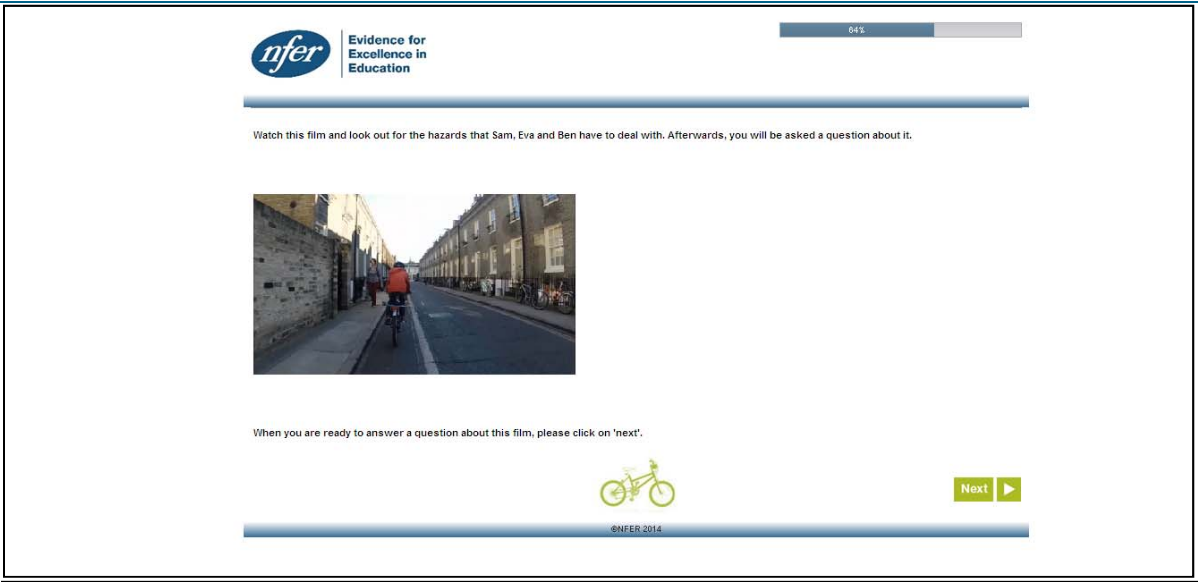
		Test 1 – Item 17					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.43						
Facility	21	14	25	26	18	22	20
Discrimination	0.21						
IRT Slope	-						
Domain	Observation						
Common item	×						
Notes	Item removed due to poor functioning.						

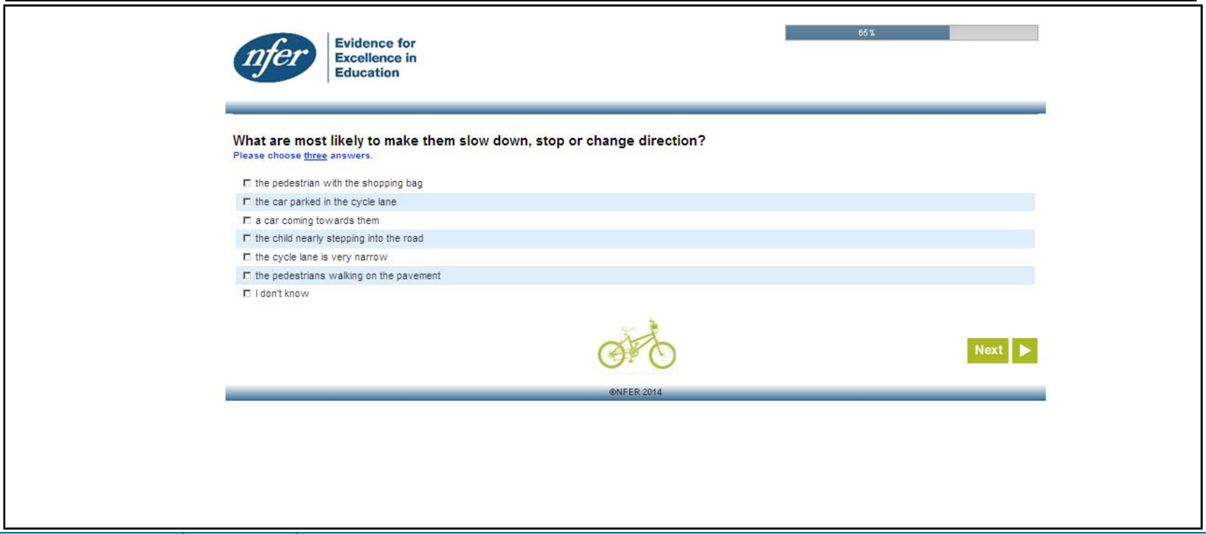


		Test 1/2 – Item 18					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	**Trained**	**Comparison**	**Trained**	**Comparison**	**Trained**	**Comparison**
Max score	2						
Mean score Test 1 Test 2	 0.85 0.81						
Facility Test 1 Test 2	 42 41	44	39	44	39	43	44
Discrimination Test 1 Test 2	 0.21 0.21						
IRT Slope	-						
Domain	Level 1						
Common item	✓						
Notes	Item removed due to poor functioning.						



		Test 1 – Item 19					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.11						
Facility	11	6	14	16	13	9	9
Discrimination	0.06						
IRT Slope	-						
Domain	Rights of way						
Common item	×						
Notes	Favours girls 5% level of significance. Item removed due to poor functioning.						

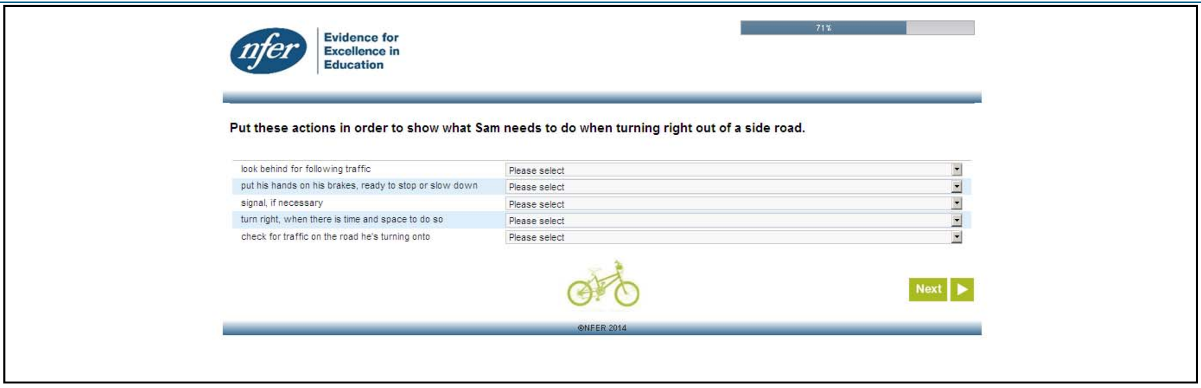




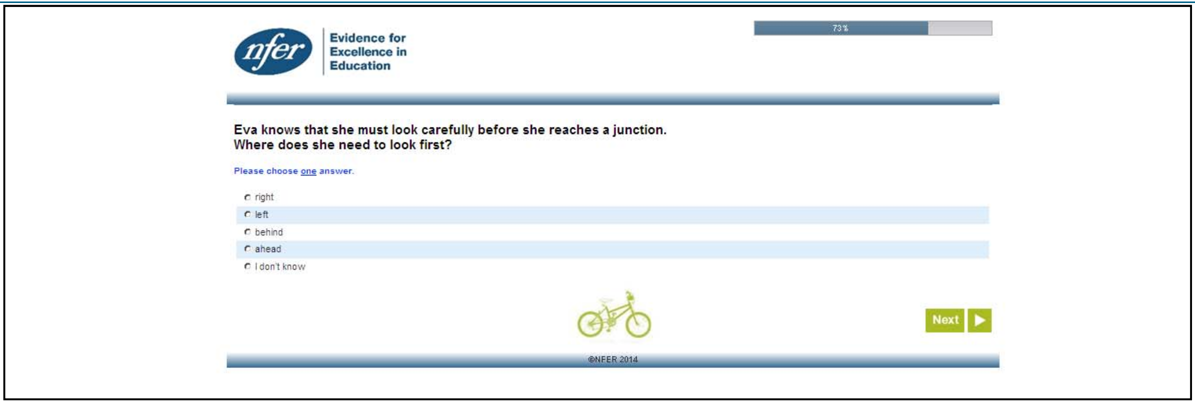
		Test 1 – Item 20					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.76						
Facility	38	35	36	30	39	48	40
Discrimination	0.22						
IRT Slope	0.31						
Domain	Observation						
Common item	✗						
Notes	Overall: 62% obtained 1 mark; 7% obtained 2 marks						

		Test 1/2 – Item 21					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	0.46 0.47						
Facility Test 1 Test 2	46 47	44	47	46	50	54	43
Discrimination Test 1 Test 2	0.23 0.15						
IRT Slope	-						
Domain	Observation						
Common item	✓						
Notes	Item removed due to poor functioning.						

		Test 1/2 – Item 22					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	0.21 0.21						
Facility Test 1 Test 2	21 21	6	13	61	23	51	10
Discrimination Test 1 Test 2	0.33 0.36						
IRT Slope	0.69						
Domain	Road position						
Common item	✓						
Notes							

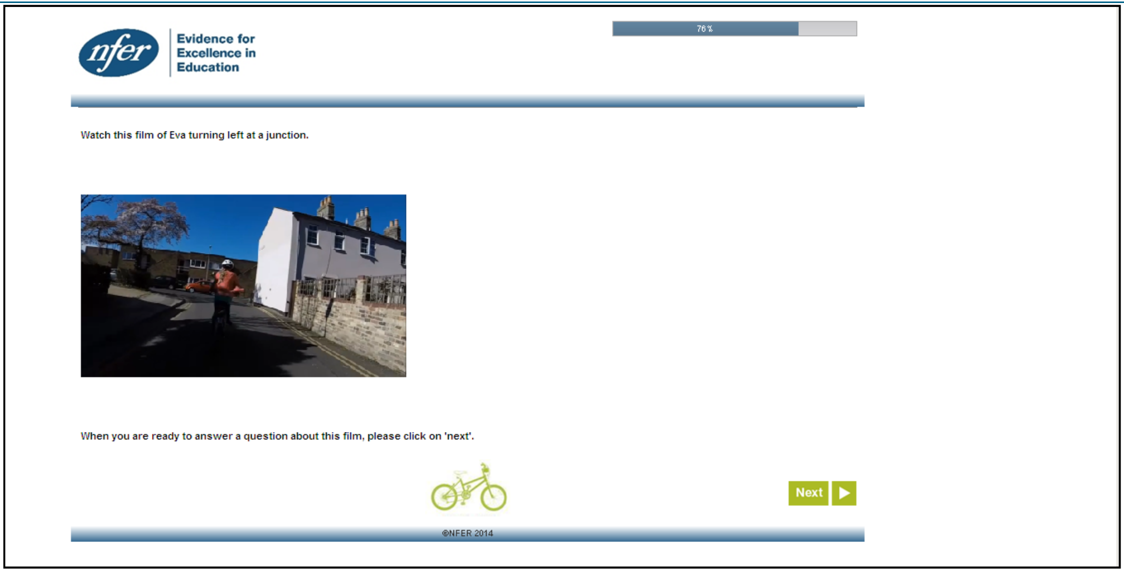


		Test 1/2 – Item 23					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	 0.07 0.08						
Facility Test 1 Test 2	 7 8	5	6	20	3	14	4
Discrimination Test 1 Test 2	 0.16 0.23						
IRT Slope	0.50						
Domain	Observation, Communication, Road Position and Priorities						
Common item	✓						
Notes							

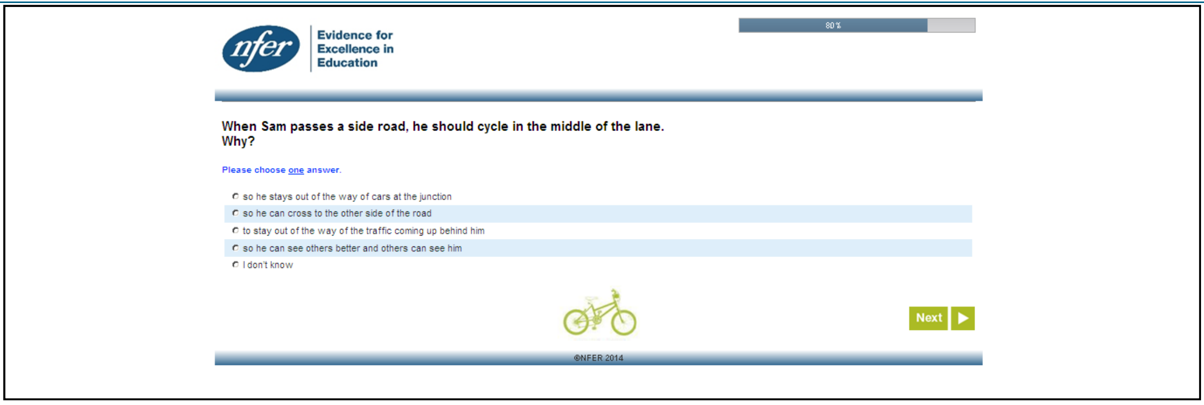


		Test 1 – Item 24					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.46						
Facility	46	29	40	68	51	65	44
Discrimination	0.35						
IRT Slope	0.57						
Domain	Observation						
Common item	×						
Notes							

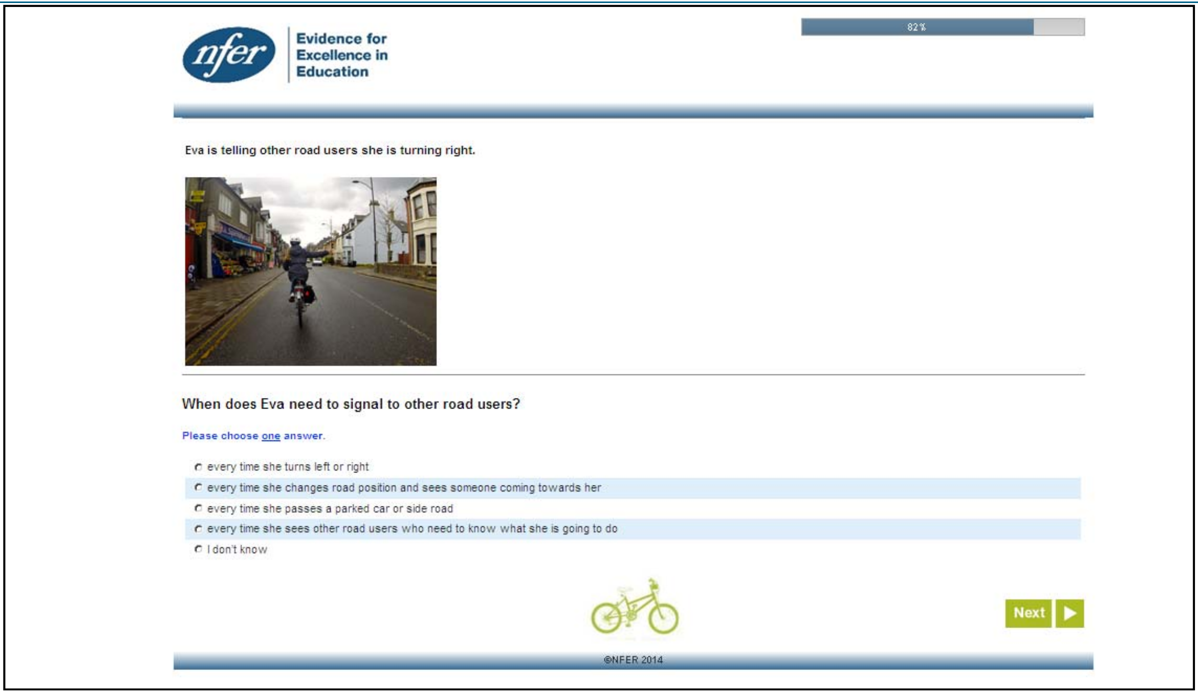
		Test 1/2 – Item 25					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	0.23 0.35						
Facility Test 1 Test 2	23 35	26	25	34	30	33	34
Discrimination Test 1 Test 2	0.14 0.15						
IRT Slope	-						
Domain	Rights of way						
Common item	✓						
Notes	Favours boys 5% level of significance. Item removed due to poor functioning.						



		Test 1/2 – Item 26					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	 0.47 0.52						
Facility Test 1 Test 2	 47 52	53	41	66	40	66	51
Discrimination Test 1 Test 2	 0.42 0.37						
IRT Slope	0.66						
Domain	Rights of way						
Common item	✓						
Notes							

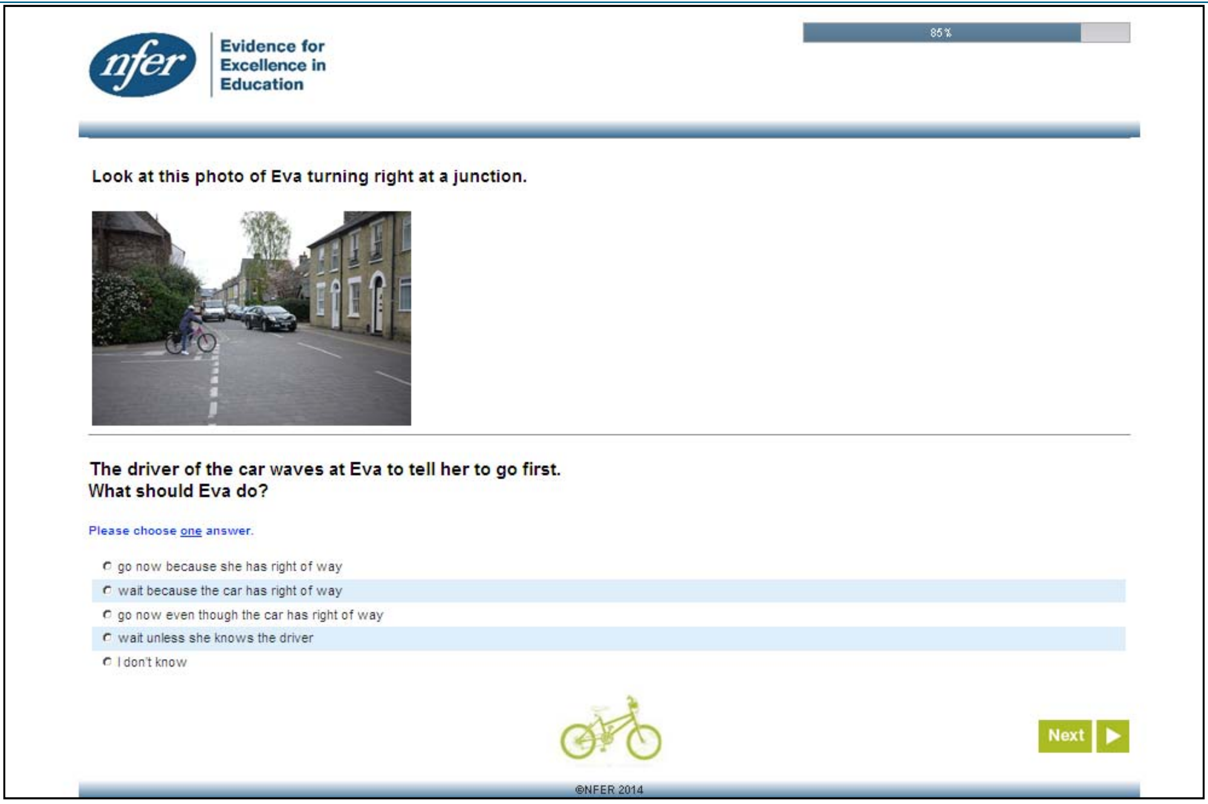


		Test 1 – Item 27					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.26						
Facility	26	31	26	26	20	29	21
Discrimination	0.19						
IRT Slope	0.29						
Domain	Road position						
Common item	×						
Notes							



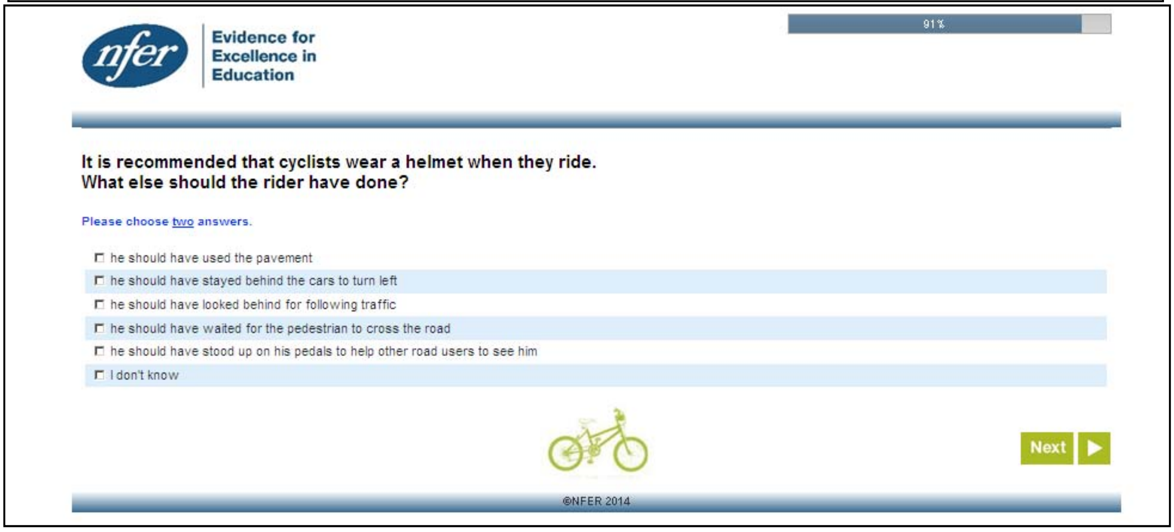
		Test 1/2 – Item 28					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	0.20 0.22						
Facility Test 1 Test 2	20 22	15	16	34	26	33	20
Discrimination Test 1 Test 2	0.25 0.23						
IRT Slope	0.42						
Domain	Communication						
Common item	✓						
Notes							

		Test 1 – Item 29						
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.69						
Facility	34	27	34	47	21	48	31
Discrimination	0.31						
IRT Slope	0.43						
Domain	Observation						
Common item	✗						
Notes	Overall: 34% obtained 1 mark; 17% obtained 2 marks						

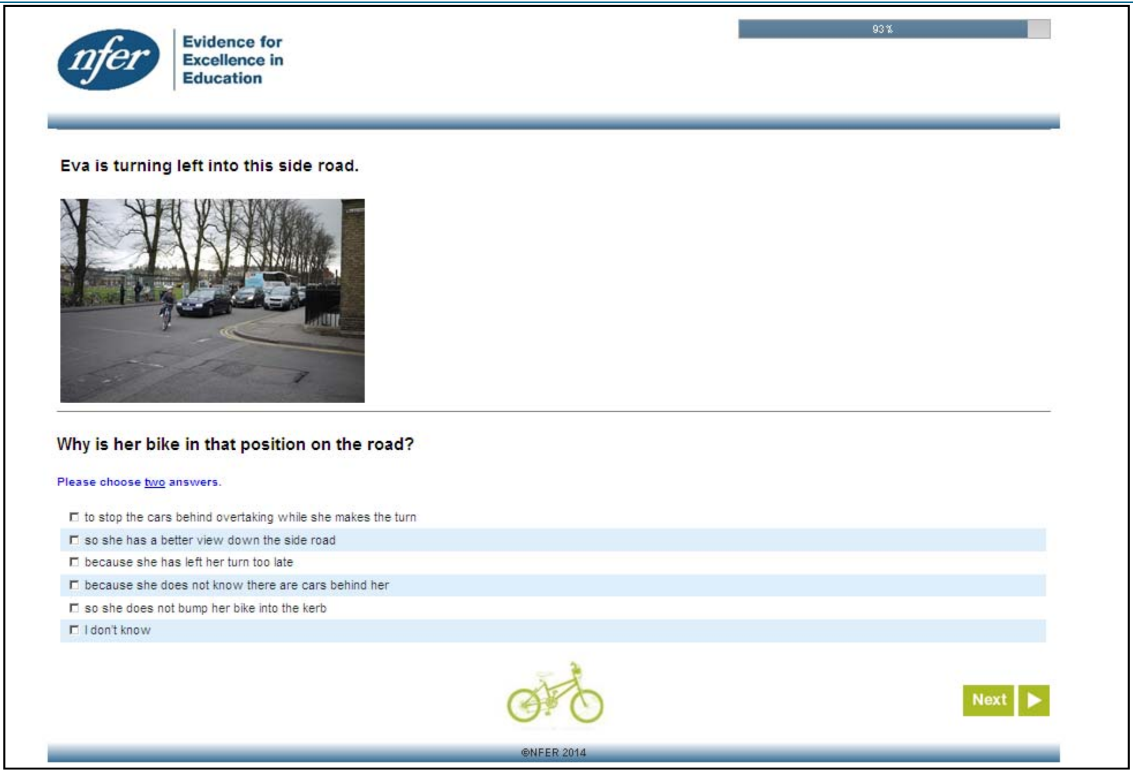


		Test 1/2 – Item 30					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score Test 1 Test 2	0.25 0.29						
Facility Test 1 Test 2	25 29	19	25	41	34	42	19
Discrimination Test 1 Test 2	0.14 0.04						
IRT Slope	-						
Domain	Rights of way						
Common item	✓						
Notes	Item removed due to poor functioning.						

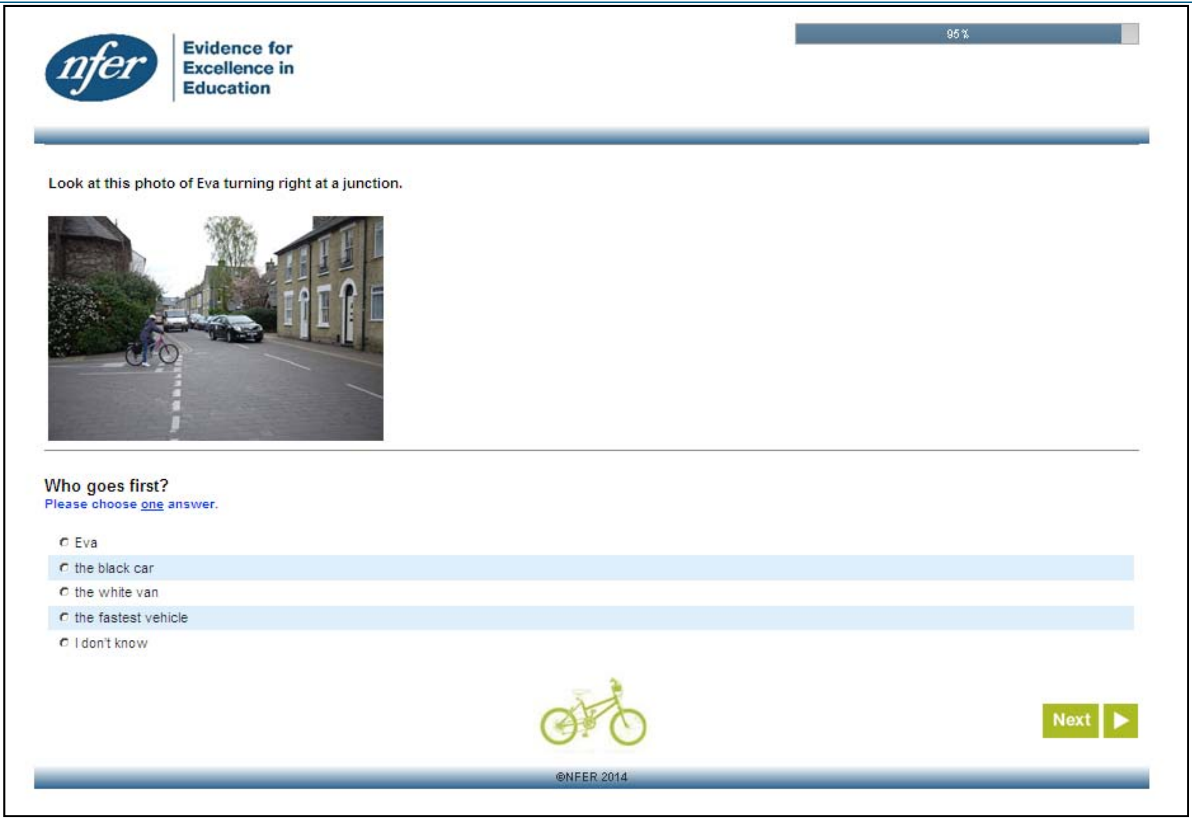
		Test 1/2 – Item 31					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score Test 1 Test 2	0.96 0.96						
Facility Test 1 Test 2	48 48	52	42	58	44	57	47
Discrimination Test 1 Test 2	0.39 0.41						
IRT Slope	0.70						
Domain	Communication						
Common item	✓						
Notes	Test 1 - overall: 76% obtained 1 mark; 10% obtained 2 marks Test 2 - overall: 73% obtained 1 mark; 11% obtained 2 marks						



		Test 1 – Item 32					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.58						
Facility	29	34	24	41	19	36	32
Discrimination	0.32						
IRT Slope	0.45						
Domain	Observation & Road position						
Common item	×						
Notes	Overall: 23% obtained 1 mark; 18% obtained 2 marks						



		Test 1/2 – Item 33					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	**Trained**	**Comparison**	**Trained**	**Comparison**	**Trained**	**Comparison**
Max score	2						
Mean score Test 1 Test 2	0.68 0.64						
Facility Test 1 Test 2	34 32	30	30	49	27	48	26
Discrimination Test 1 Test 2	0.26 0.30						
IRT Slope	0.35						
Domain	Road position						
Common item	✓						
Notes	Test 1 - overall: 27% obtained 1 mark; 20% obtained 2 marks Test 2 - overall: 30% obtained 1 mark; 17% obtained 2 marks						



		Test 1 – Item 34					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.49						
Facility	49	40	53	58	38	38	57
Discrimination	0.09						
IRT Slope	-						
Domain	Rights of way						
Common item	×						
Notes	Item removed due to poor functioning.						

		Test 1 – Item 35					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.36						
Facility	36	13	32	66	36	65	26
Discrimination	0.30						
IRT Slope	0.48						
Domain	Observation						
Common item	✗						
Notes							

Evidence for Excellence in Education

22%

Look at Sam's bike.

Which parts of his bike should Sam check before every ride?

Please choose three answers.

- pedals
- brakes
- saddle
- chain
- air
- wheels
- handlebars
- frame

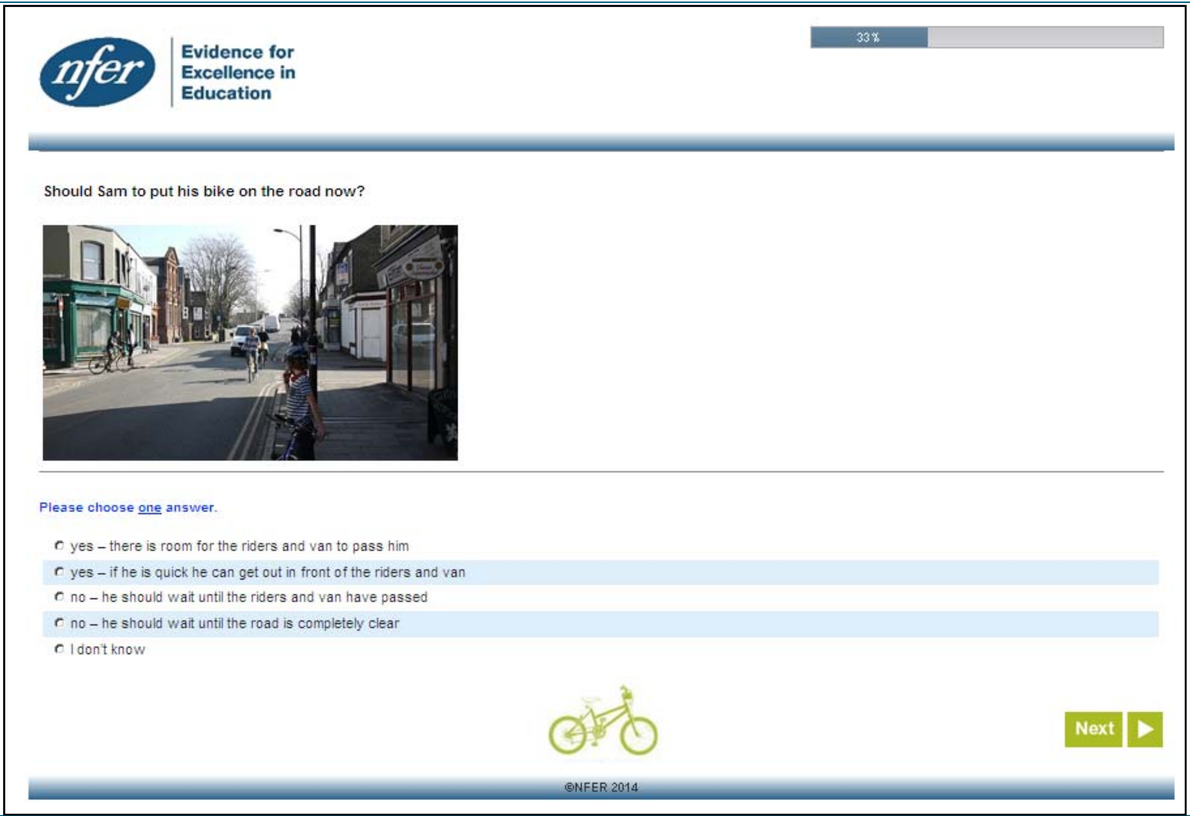
Next

©NFER 2014

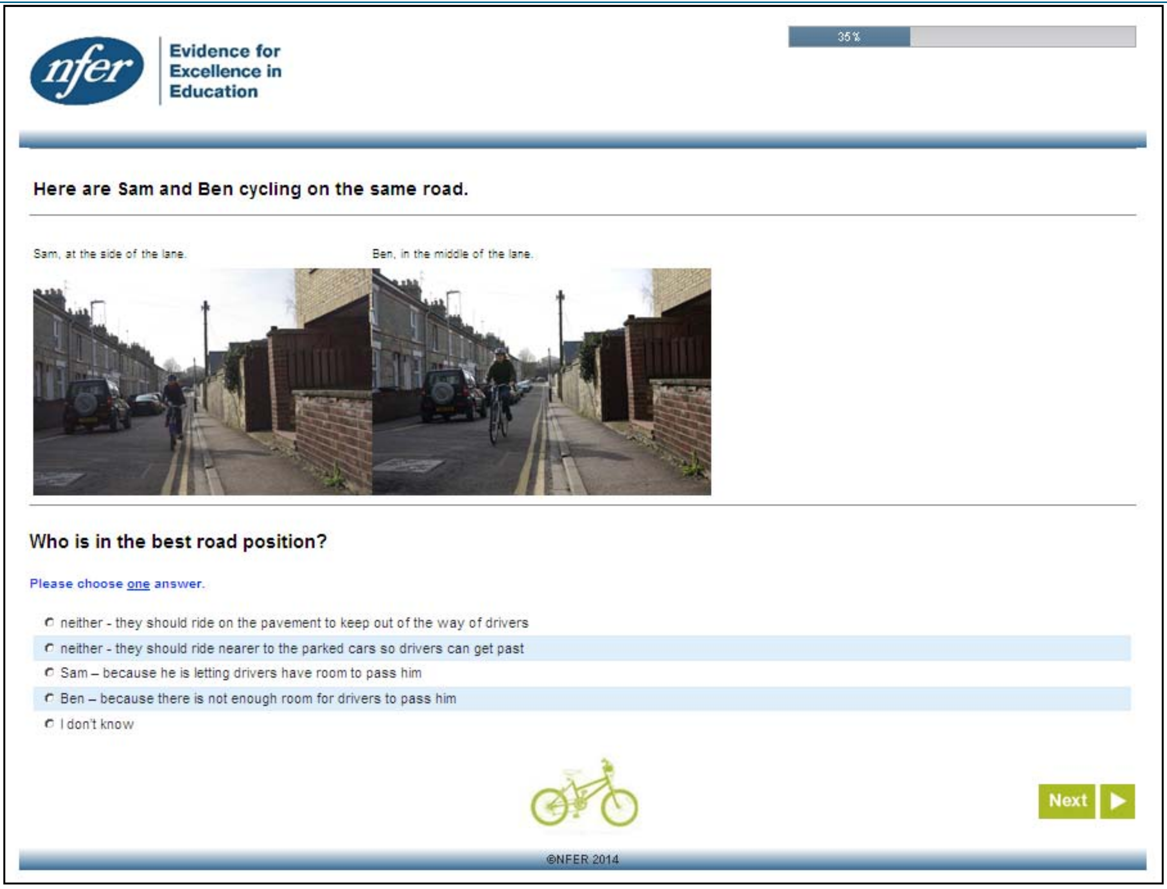
		Test 2 – Item 36					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.99						
Facility	50	50	45	66	38	61	50
Discrimination	0.26						
IRT Slope	0.36						
Domain	Level 1						
Common item	✗						
Notes	Overall: 55% obtained 1 mark; 22% obtained 2 marks						

		Test 2 – Item 37					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	**Trained**	**Comparison**	**Trained**	**Comparison**	**Trained**	**Comparison**
Max score	1						
Mean score	0.51						
Facility	51	57	50	34	54	50	56
Discrimination	0.14						
IRT Slope	-						
Domain	Road position						
Common item	×						
Notes	Item removed due to poor functioning.						

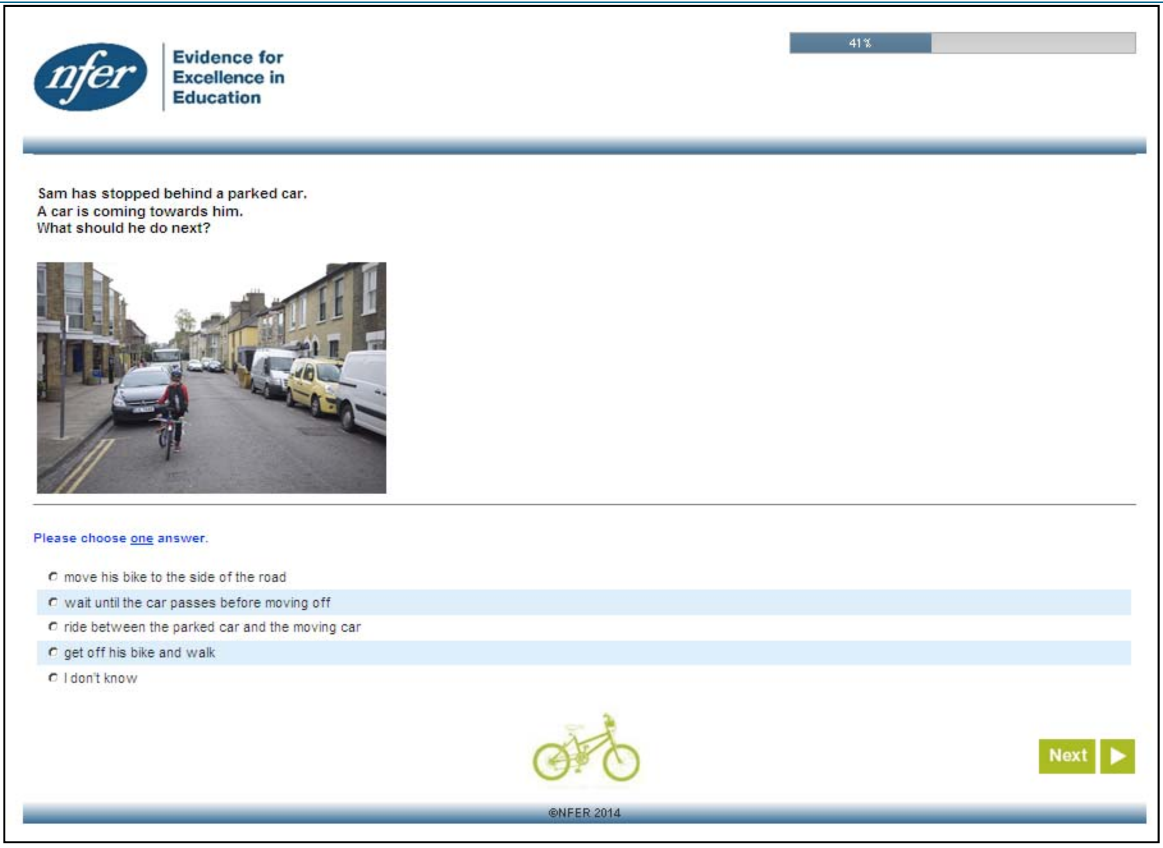
		Test 2 – Item 38					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.54						
Facility	54	44	41	73	61	90	55
Discrimination	0.38						
IRT Slope	0.68						
Domain	Observation						
Common item	×						
Notes							



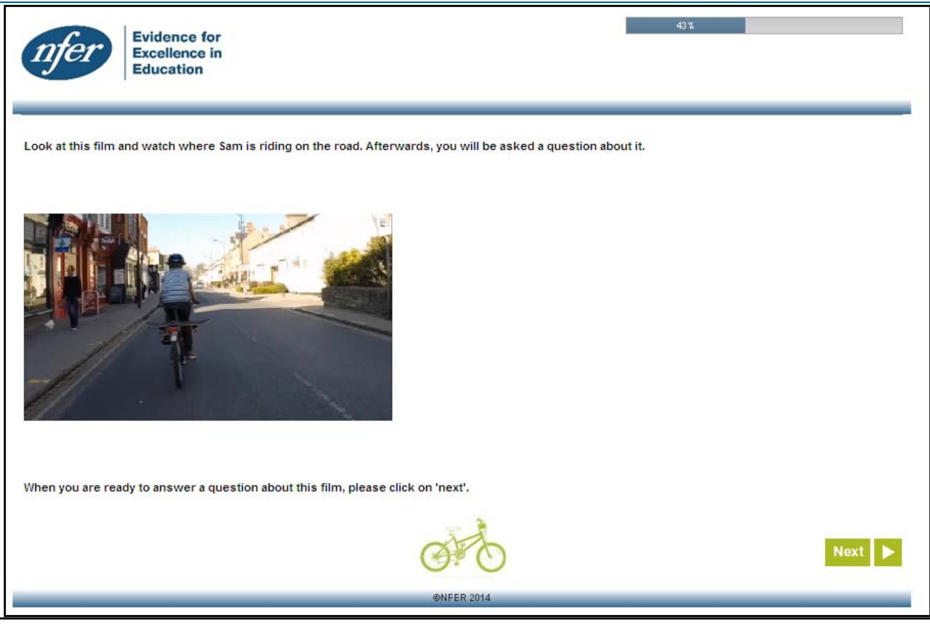
		Test 2 – Item 39					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.52						
Facility	52	46	49	66	52	63	51
Discrimination	0.15						
IRT Slope	-						
Domain	Observation						
Common item	×						
Notes	Item removed due to poor functioning.						



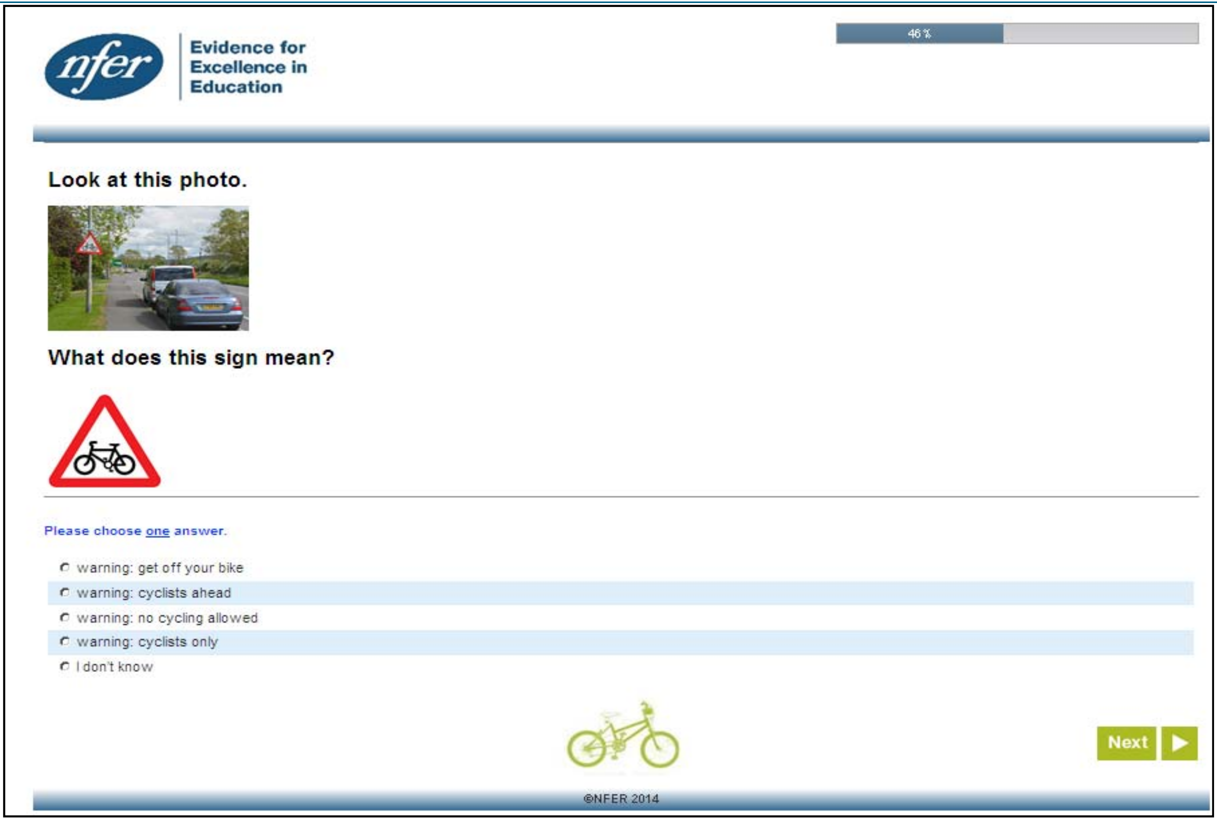
		Test 2 – Item 40					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.15						
Facility	15	3	11	46	20	35	3
Discrimination	0.28						
IRT Slope	0.60						
Domain	Road position						
Common item	×						
Notes							



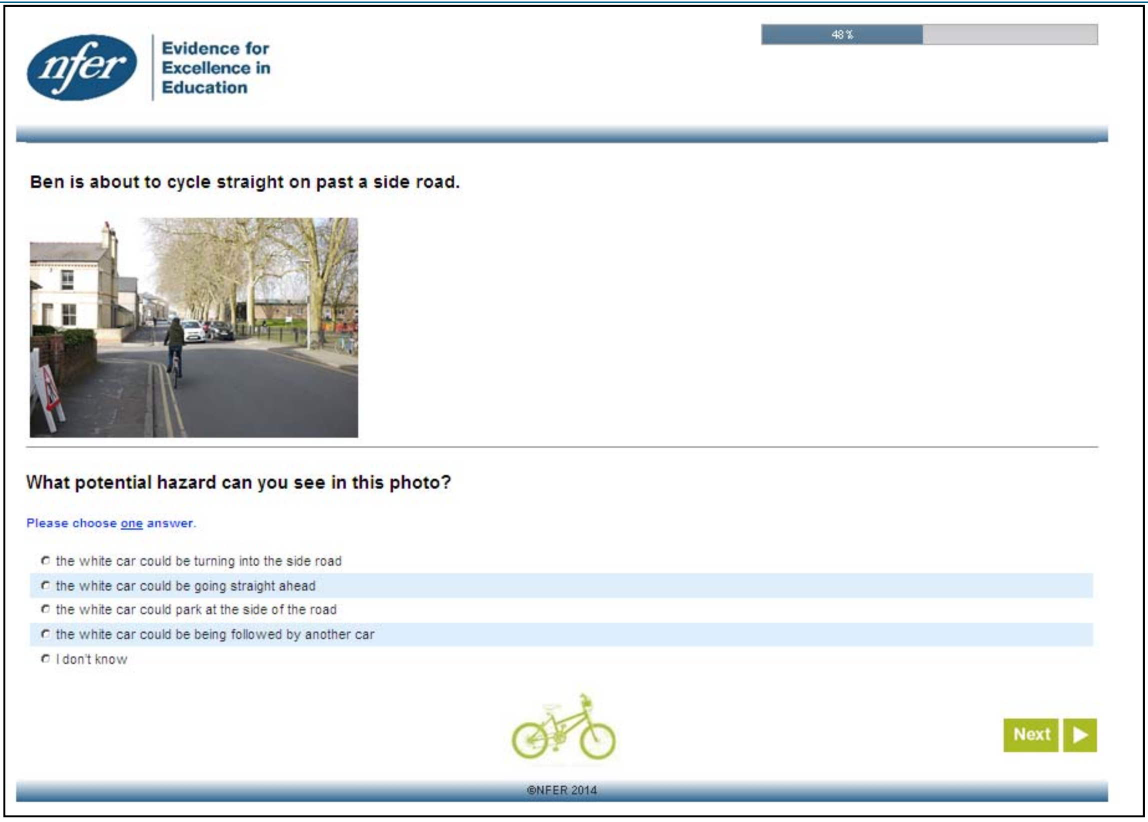
		Test 2 – Item 41					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.57						
Facility	57	54	46	73	63	78	60
Discrimination	0.35						
IRT Slope	0.57						
Domain	Rights of way						
Common item	✗						
Notes							



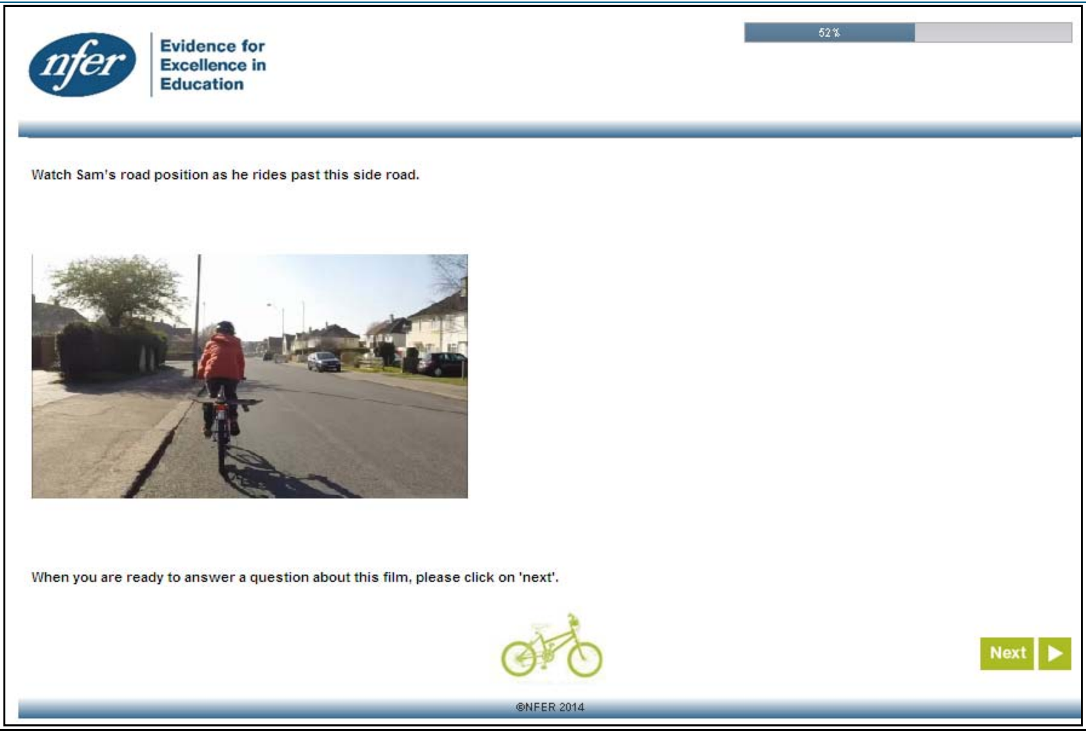
		Test 2 – Item 42					
		Phase 1 (Baseline)		**Phase 2 (Post-training - June)**		**Phase 3 (Post-training – September)**	
	Overall	**Trained**	**Comparison**	**Trained**	**Comparison**	**Trained**	**Comparison**
Max score	2						
Mean score	0.57						
Facility	28	23	23	48	36	51	19
Discrimination	0.29						
IRT Slope	0.41						
Domain	Road position						
Common item	✗						
Notes	Overall: 34% obtained 1 mark; 11% obtained 2 marks						

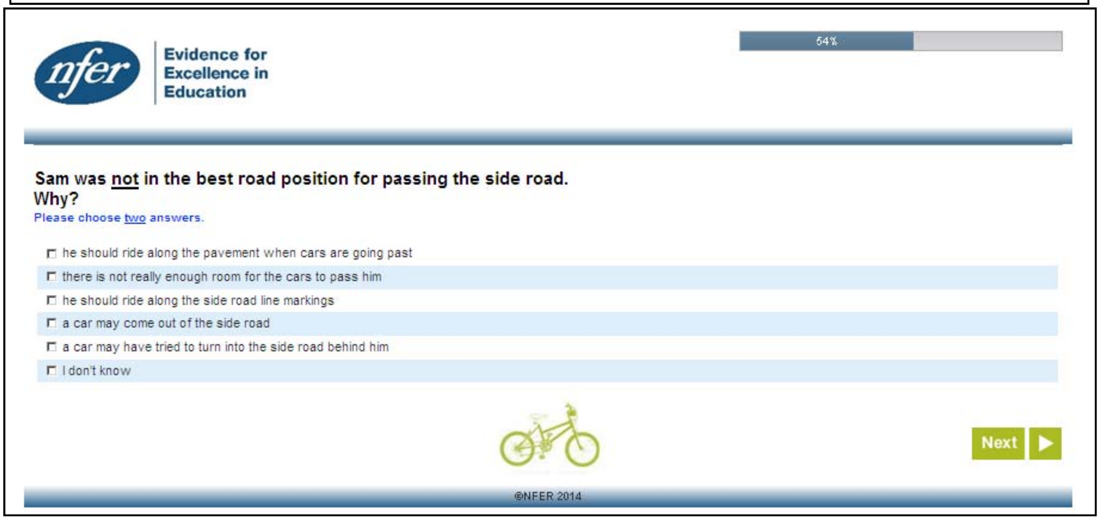


		Test 2 – Item 43					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.53						
Facility	53	57	53	51	39	43	61
Discrimination	0.09						
IRT Slope	-						
Domain	Rights of way						
Common item	×						
Notes	Item removed due to poor functioning.						

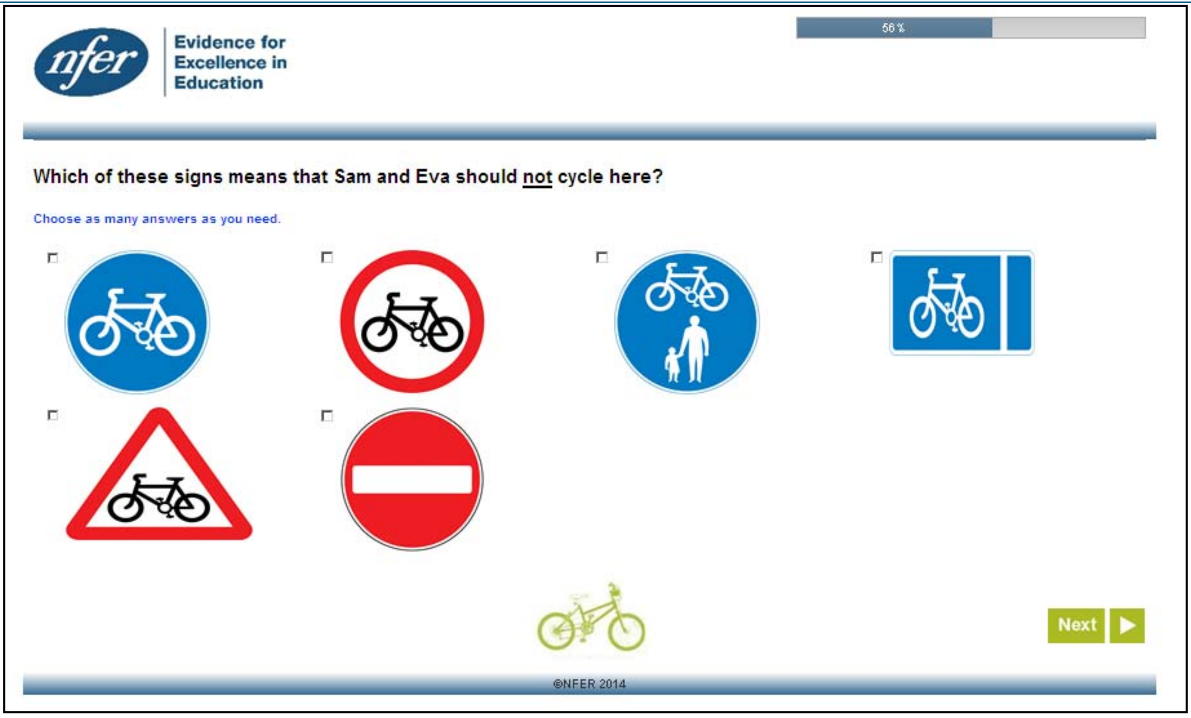


		Test 2 – Item 44					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.73						
Facility	73	79	70	63	67	82	78
Discrimination	0.31						
IRT Slope	0.54						
Domain	Observation						
Common item	×						
Notes							

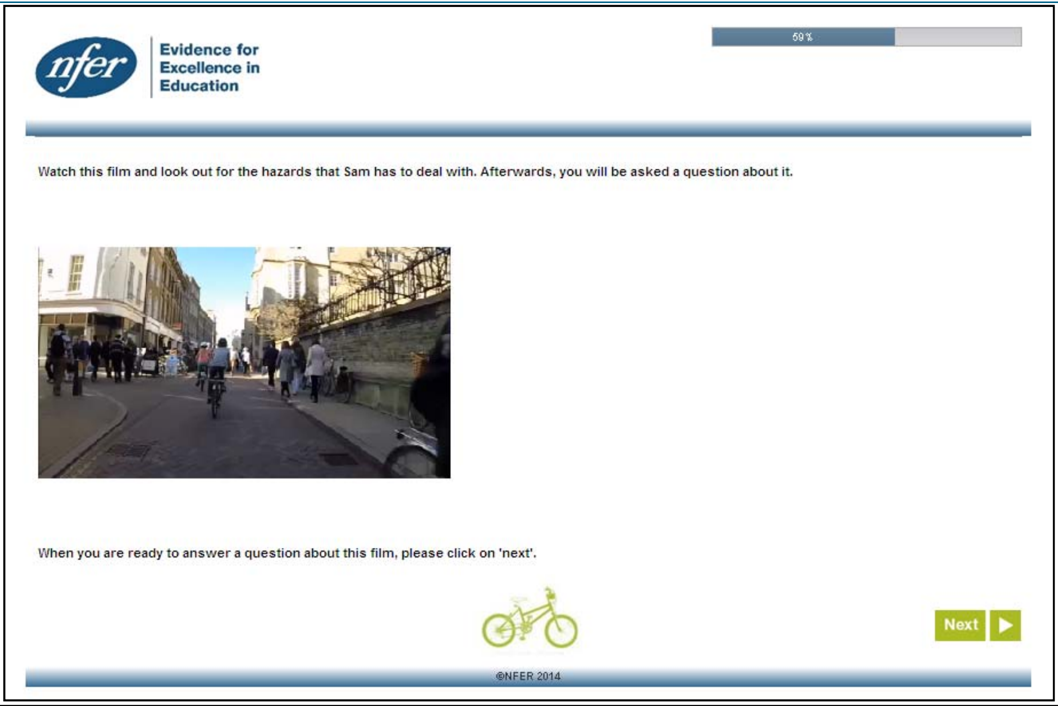


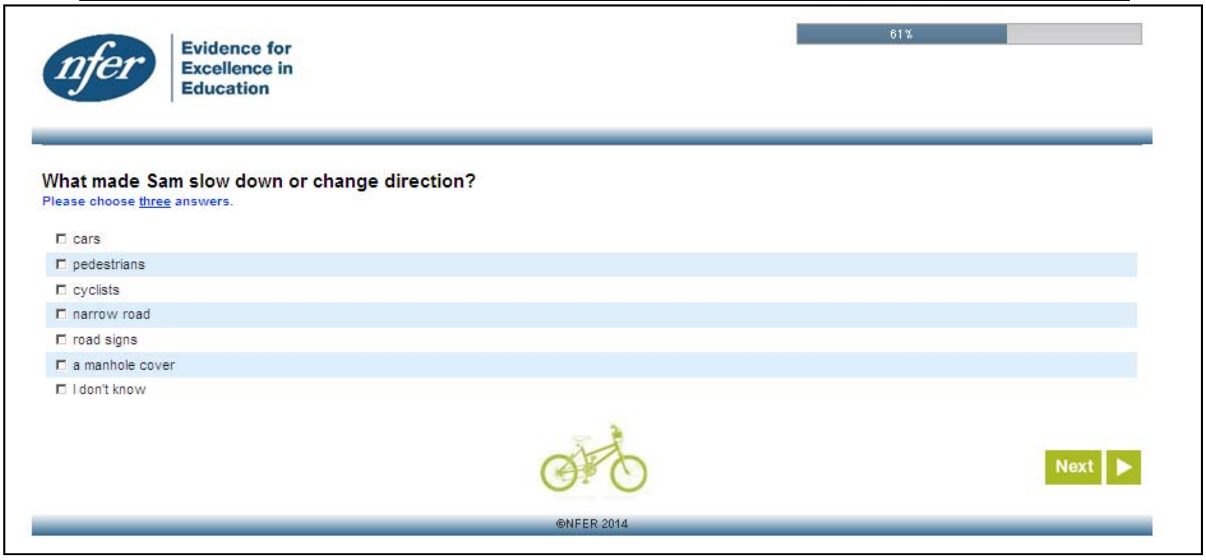


		Test 2 – Item 45					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.38						
Facility	19	21	16	22	17	28	17
Discrimination	0.20						
IRT Slope	0.34						
Domain	Road position						
Common item	✗						
Notes	Overall: 22% obtained 1 mark; 8% obtained 2 marks						



		Test 2 – Item 46					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.58						
Facility	29	26	27	29	21	38	32
Discrimination	0.19						
IRT Slope	0.28						
Domain	Rights of way						
Common item	×						
Notes	Overall: 46% obtained 1 mark; 6% obtained 2 marks						





		Test 2 – Item 47					
		Phase 1 (Baseline)		**Phase 2 (Post-training - June)**		**Phase 3 (Post-training – September)**	
	Overall	**Trained**	**Comparison**	**Trained**	**Comparison**	**Trained**	**Comparison**
Max score	2						
Mean score	1.01						
Facility	51	67	43	45	37	64	55
Discrimination	0.32						
IRT Slope	0.45						
Domain	Observation						
Common item	✗						
Notes	Overall: 53% obtained 1 mark; 24% obtained 2 marks						

		Test 2 – Item 48					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.61						
Facility	61	61	58	59	57	80	60
Discrimination	0.32						
IRT Slope	0.49						
Domain	Rights of way						
Common item	×						
Notes							

		Test 2 – Item 49					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.63						
Facility	63	67	59	63	57	80	61
Discrimination	0.21						
IRT Slope	0.31						
Domain	Observation						
Common item	×						
Notes							

Eva is approaching a junction to turn left.
She looks behind and sees a car following her.

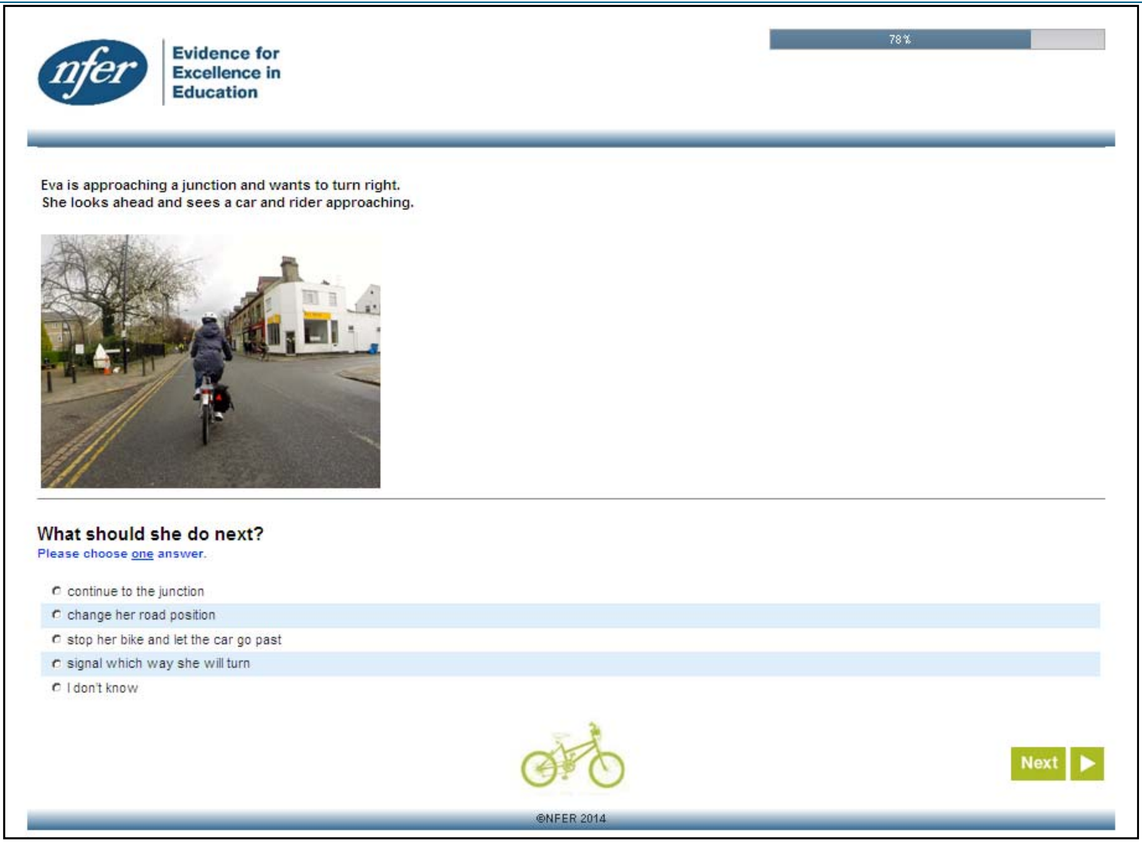
What should she do next?
Please choose one answer.

- change her road position
- stop her bike and let the car go past
- signal which way she will turn
- race the car to the junction
- I don't know

Next

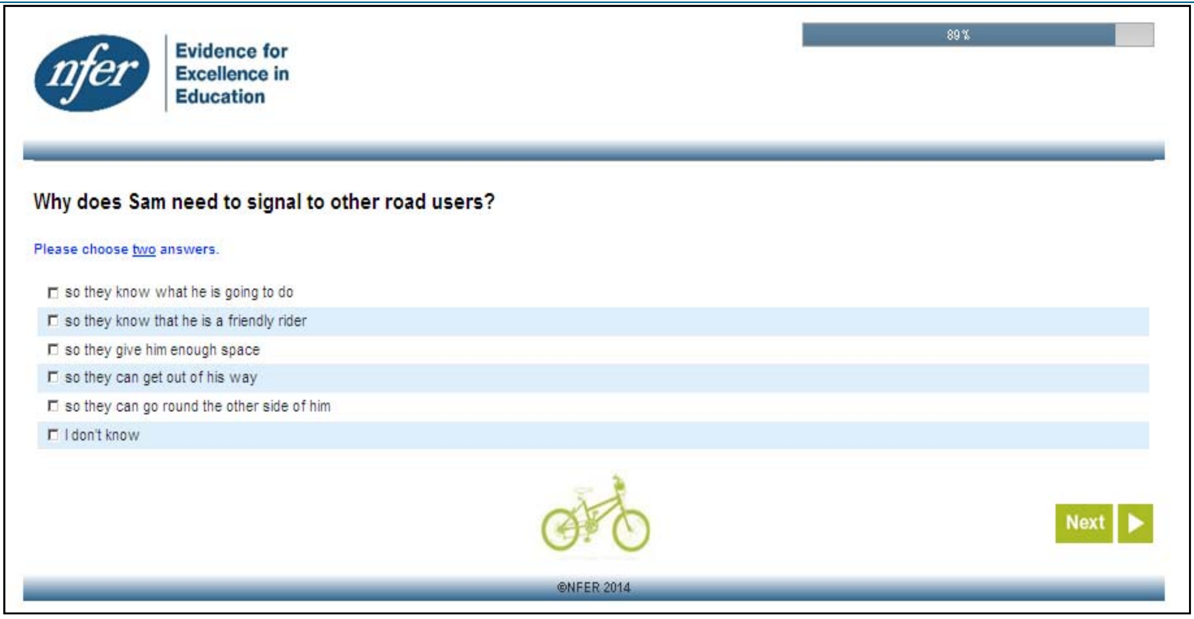
©NFER 2014

		Test 2 – Item 50					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.65						
Facility	65	80	55	61	52	87	69
Discrimination	0.44						
IRT Slope	0.84						
Domain	Communication						
Common item	×						
Notes							

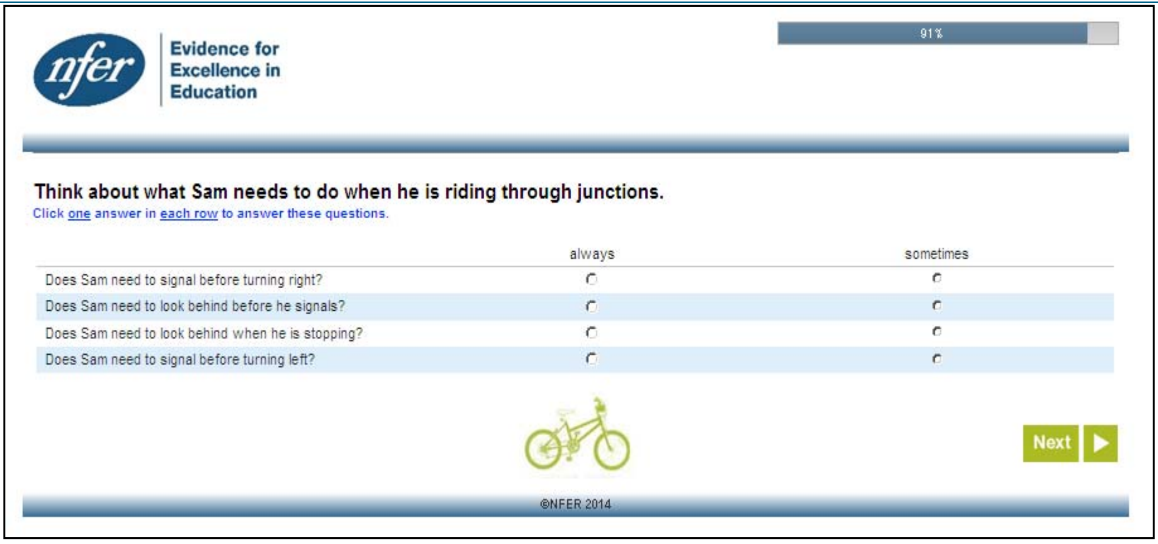


		Test 2 – Item 51					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.51						
Facility	51	63	48	34	37	65	54
Discrimination	0.26						
IRT Slope	0.38						
Domain	Communication						
Common item	×						
Notes							

		Test 2 – Item 52					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.12						
Facility	12	7	12	20	9	15	10
Discrimination	-0.02						
IRT Slope	-						
Domain	Observation						
Common item	×						
Notes	Item removed due to poor functioning.						



		Test 2 – Item 53					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.96						
Facility	48	54	43	48	52	61	46
Discrimination	0.25						
IRT Slope	0.32						
Domain	Communication						
Common item	×						
Notes	Overall: 34% obtained 1 mark; 32% obtained 2 marks						



		Test 2 – Item 54					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	2						
Mean score	0.45						
Facility	22	11	19	44	28	34	18
Discrimination	0.29						
IRT Slope	0.46						
Domain	Observation & Communication						
Common item	✗						
Notes	Overall: 8% obtained 1 mark; 18% obtained 2 marks						

		Test 2 – Item 55					
		Phase 1 (Baseline)		Phase 2 (Post-training - June)		Phase 3 (Post-training – September)	
	Overall	Trained	Comparison	Trained	Comparison	Trained	Comparison
Max score	1						
Mean score	0.31						
Facility	31	23	19	42	28	58	18
Discrimination	0.28						
IRT Slope	0.43						
Domain	Observation						
Common item	×						
Notes							

E.2 Further analysis

The following tables give a more detailed presentation of the statistical analysis conducted to compare the outcome differences between the trained and comparison groups. The tables present the coefficients from the multilevel models, as well as the standard errors and 95 per cent confidence intervals of the coefficients. In addition, the intra-cluster correlation is presented, which shows how much variation is between schools rather than between pupils.

The 'pupil-level standard deviation' is the pupil-level standard deviation at phase 1 from a multilevel model with no covariates. The effect size is defined as the coefficient on the trained group indicator variable divided by the pupil-level standard deviation.

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Hazard perception and appropriate response ability (phase 2)				
Intercept	5.63	8.16	-10.4 – 21.6	0.49
Trained group	28.30	2.89	22.6 – 34.0	0.00
Hazard perception and appropriate response ability (phase 1)	0.96	0.08	0.8 – 1.1	0.00
Number of pupils	142			
Intra-cluster correlation	0.17			
Pupil-level standard deviation	17.91			
Standardised effect size	1.58	0.16		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Hazard perception and appropriate response ability (phase 3)				
Intercept	28.82	9.66	9.9 – 47.8	0.00
Trained group	28.72	3.29	22.3 – 35.2	0.00
Hazard perception and appropriate response ability (phase 1)	0.71	0.09	0.5 – 0.9	0.00
Number of pupils	154			
Intra-cluster correlation	0.17			
Pupil-level standard deviation (no covariates)	17.91			
Standardised effect size	1.60	0.18		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Observation sub-domain (phase 2)				
Intercept	0.23	0.05	0.1 – 0.3	0.00
Trained group	0.17	0.06	0.1 – 0.3	0.03
Observation sub-domain (phase 1)	0.45	0.09	0.3 – 0.6	0.00
Number of pupils	142			
Intra-cluster correlation	0.10			
Pupil-level standard deviation	0.16			
Standardised effect size	1.05	0.34		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Observation sub-domain (phase 3)				
Intercept	0.18	0.04	0.1 – 0.3	0.00
Trained group	0.17	0.03	0.1 – 0.2	0.00
Observation sub-domain (phase 1)	0.60	0.10	0.4 – 0.8	0.00
Number of pupils	154			
Intra-cluster correlation	0.10			
Pupil-level standard deviation	0.16			
Standardised effect size	1.07	0.18		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Communication sub-domain (phase 2)				
Intercept	0.11	0.04	0.0 – 0.2	0.02
Trained group	0.13	0.03	0.1 – 0.2	0.01
Communication sub-domain (phase 1)	0.69	0.11	0.5 – 0.9	0.00
Number of pupils	142			
Intra-cluster correlation	0.07			
Pupil-level standard deviation	0.17			
Standardised effect size	0.77	0.20		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Communication sub-domain (phase 3)				
Intercept	0.18	0.05	0.1 – 0.3	0.00
Trained group	0.13	0.04	0.0 – 0.2	0.02
Communication sub-domain (phase 1)	0.46	0.11	0.3 – 0.7	0.00
Number of pupils	154			
Intra-cluster correlation	0.07			
Pupil-level standard deviation	0.17			
Standardised effect size	0.75	0.25		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Road position sub-domain (phase 2)				
Intercept	0.22	0.05	0.1 – 0.3	0.00
Trained group	0.22	0.06	0.1 – 0.3	0.02
Road position sub-domain (phase 1)	0.23	0.11	0.0 – 0.4	0.04
Number of pupils	142			
Intra-cluster correlation	0.06			
Pupil-level standard deviation	0.16			
Standardised effect size	1.39	0.40		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Road position sub-domain (phase 3)				
Intercept	0.15	0.04	0.1 – 0.2	0.00
Trained group	0.24	0.04	0.2 – 0.3	0.00
Road position sub-domain (phase 1)	0.14	0.11	-0.1 – 0.3	0.21
Number of pupils	154			
Intra-cluster correlation	0.06			
Pupil-level standard deviation	0.16			
Standardised effect size	1.54	0.27		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Priorities sub-domain (phase 2)				
Intercept	0.16	0.04	0.1 – 0.2	0.00
Trained group	0.24	0.04	0.2 – 0.3	0.00
Priorities sub-domain (phase 1)	0.41	0.09	0.2 – 0.6	0.00
Number of pupils	142			
Intra-cluster correlation	0.10			
Pupil-level standard deviation	0.20			
Standardised effect size	1.21	0.18		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Road position sub-domain (phase 3)				
Intercept	0.26	0.04	0.2 – 0.3	0.00
Trained group	0.20	0.03	0.1 – 0.3	0.00
Road position sub-domain (phase 1)	0.15	0.09	0.0 – 0.3	0.09
Number of pupils	154			
Intra-cluster correlation	0.10			
Pupil-level standard deviation	0.20			
Standardised effect size	1.02	0.18		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Cycling confidence (phase 3)				
Intercept	1.90	0.22	1.5 – 2.3	0.00
Trained group	0.47	0.11	0.3 – 0.7	0.00
Cycling confidence (phase 1)	0.37	0.07	0.2 – 0.5	0.00
Number of pupils	152			
Intra-cluster correlation	0.07			
Pupil-level standard deviation	0.88			
Standardised effect size	0.53	0.12		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Cycling enjoyment (phase 3)				
Intercept	2.01	0.32	1.4 – 2.6	0.00
Trained group	0.19	0.12	-0.1 – 0.4	0.16
Cycling enjoyment (phase 1)	0.52	0.07	0.4 – 0.7	0.00
Number of pupils	141			
Intra-cluster correlation	0.02			
Pupil-level standard deviation	0.96			
Standardised effect size	0.20	0.13		

Variable	Coefficient	Standard error	Confidence interval	p-value
Dependent = Cycling frequency (phase 3)				
Intercept	1.13	0.27	0.6 – 1.7	0.00
Trained group	-0.02	0.19	-0.4 – 0.4	0.91
Cycling frequency (phase 1)	0.62	0.07	0.5 – 0.8	0.00
Number of pupils	149			
Intra-cluster correlation	0.04			
Pupil-level standard deviation (no covariates)	1.35			
Standardised effect size	-0.02	0.14		

P-values relating to the correlation between on-screen quiz and practical assessment scores (Table 3.9)

	Phase 2		Phase 3	
Domain	**Correlation**	**p-value**	**Correlation**	**p-value**
Observation	0.48	0.00	0.28	0.03
Communication	0.28	0.06	0.29	0.02
Road position	0.22	0.14	0.24	0.06
Priorities	0.40	0.01	0.02	0.91
Overall	0.40	0.01	0.35	0.01
Number of pupils	*48*		*60*	

E.3 Statistics glossary

Correlation

The correlation between two variables describes the extent to which they vary together. Two variables have a high correlation where if one is high the other is consistently high (or low if negatively correlated) and if it is low then the other is consistently low (or high if negatively correlated).

The correlation coefficient is a measure of correlation that varies between -1 and +1, where -1 indicates perfect negative correlation (when one is high the other is low, and vice versa) and where +1 indicates perfect positive correlation (when one is high the other is high, and vice versa).

Confidence intervals

Confidence intervals provide the range of values that has a 95 per cent probability of including the true effect size. The width of the confidence interval indicates the confidence we can place in a finding: the wider the interval, the less confidence we can have. If we repeated the research 100 times we would expect different answers each time, but the 95 per cent confidence interval gives a range within which we would expect the true answer to be, in around 95 of the confidence intervals.

If the confidence interval includes a certain value (for example, if it includes zero) then we cannot be confident that the true value is different from that and therefore it is likely that our estimate is simply down to chance.

Discrimination index

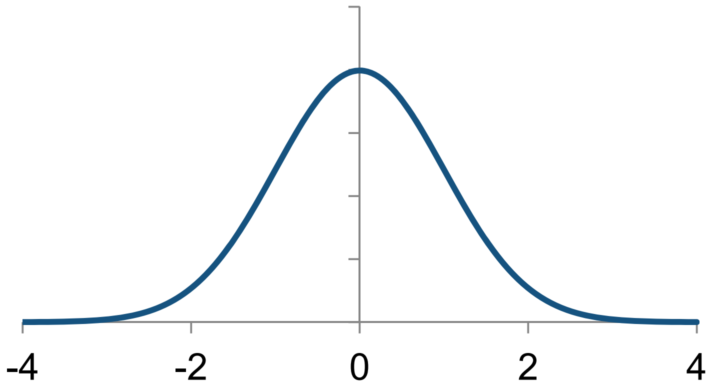
The *discrimination index* gives information about how well an item discriminates between those with high and low hazard perception and appropriate response ability. A value of 0.30 or above is normally taken as acceptable. In classical test theory (CTT) the discrimination index is most informative about items which have a facility of about 50 per cent; it is less informative about notably easy or hard items. As a rule of thumb, effective items are expected to have discrimination measures of 0.3 to 0.4 or above.

Effect size

Effect size is a way of quantifying the size of the difference between two groups. It can be applied to any measured outcome in education or social science. It is particularly valuable for quantifying the effectiveness of a particular intervention, relative to some comparison, because it is standardised. It is calculated by dividing the difference between the scores for the intervention group and the control group by the standard deviation of the underlying variable. Formally:

Effect size =

$$\frac{\text{Average outcome in intervention group - Average outcome in control group}}{\text{Standard deviation of outcome variable}}$$



This is a normal distribution curve.

Normal distribution
Mean = 0
Standard deviation = 1

Above 0 = 50%
-1 to +1 = 68%
-2 to +2 = 95%

Test scores typically have a distribution that takes roughly the shape of the normal distribution (pictured above). In a normal distribution 50 per cent of pupils are above the average and 50 per cent are below the average. The standard deviation measures the typical spread of the data: in a normal distribution 68 per cent of pupils are between one standard deviation of the average and 95 per cent of pupils are within 2 standard deviations.

Above 0 = 50%
Above 0 = 84%

An effect size is a measure of how much the outcome measure has changed, given how much it varied to start with. It is measured in standard deviations, which is the measure of spread. Pictured above is an effect size of 1: the effect is to increase the proportion of pupils that were above the average from 50 per cent to 84 per cent .

A guide to effect sizes and their associated description is provided below.

Effect size from …	… to	Description
-0.01	0.01	Very low or no effect
0.02	0.18	Low
0.19	0.44	Moderate
0.45	0.69	High
0.70	>1.0	Very high

While the effect size is useful for comparing between different interventions (e.g. 5 marks on one test might be different to 5 marks on another test) it is always important to understand the effect size in the context of the intervention. Effect sizes for educational interventions – e.g. a new way of teaching reading or maths – are usually relatively low, at around 0.2 at best, because the underlying level of knowledge is quite high. However, the literature on the impact of cycle training on hazard perception and other cycling outcomes tends to show larger effect sizes because the existing knowledge among school children is relatively low.

Facility

The *facility* of a dichotomous item is defined as the percentage of the sample attempting the test who achieve the maximum score of 1 on the item (i.e. get it right). In defining the facility of an item all missing responses are treated as incorrect thus ensuring that the percentage is of everyone attempting the test.

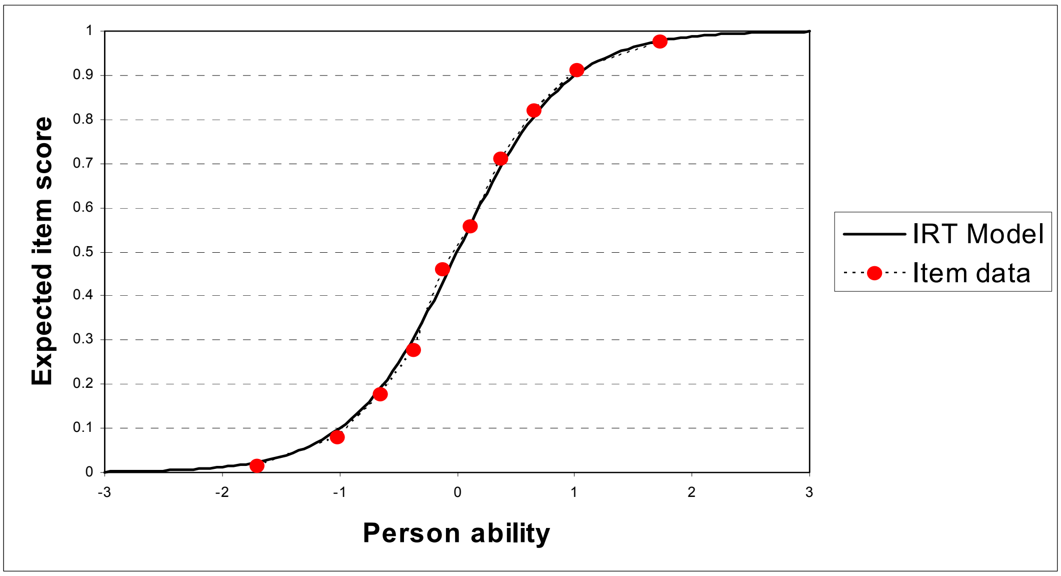
The facility of a multiple mark item is calculated here as the average score on the item as a percentage of the maximum item score. Missing responses are again treated as incorrect.

Item Response Theory (IRT)

Item Response Theory (IRT) models the relationships between pupils' ability and a set of test questions (items). IRT looks at the relationship between test and item scores based on assumptions concerning the mathematical relationship between abilities and item responses – it models the response of a pupil's given ability to each item in the test. The advantages of IRT include being able to:

- calculate how a test will work for different groups of items
- calibrate scores on different tests of different difficulty against one another.

In the following figure, the horizontal axis represents pupils' ability and the vertical axis represents the probability of a correct response to one test item. The s-shaped curve, then, shows the probabilities of a correct response for students with different ability levels.



The slope of the curve helps to determine the usefulness of the item – if a slope is less than 0.4, it is considered ‘uninformative’, i.e. the item does not tell us much about the pupils’ ability. A slope of 0.4-0.6 is classified as ‘acceptable’, 0.6-0.8 as ‘fair’, 0.8-1.2 as ‘good’ and more than 1.2 as ‘very good’. The item represented above has a slope of 1.3.

Multilevel regression modelling

Multilevel modelling is a development of a common statistical technique known as ‘regression analysis’. It explores the relationship between a measure of interest (‘dependent variable’) and the values of one or more related measures. For example, we may wish to predict average test performance given some background factors, such as performance on a previous test or pupil characteristics (sometimes called ‘independent variables’).

Multilevel modelling takes account of data which is grouped into similar clusters at different levels, such as individual pupils grouped within schools. Incorporating this hierarchical structure into our analysis improves the accuracy of its findings, and avoids drawing false or misleading conclusions from the data.

Statistical significance

Statistical significance is a test of whether or not an effect is likely to be due simply to chance. Significance is conventionally tested at the 5 per cent level of confidence: if we repeated the research 100 times we would expect different answers each time, but the 95 per cent confidence interval gives a range within which we would expect the true answer to be, in around 95 of the confidence intervals. Conversely, we would expect the true answer to be outside the 95 confidence interval in around 5 out of the 100 confidence intervals. If the confidence interval includes a certain value (for example, if it includes zero) then we cannot be confident that the true value is different from that and therefore it is likely that our estimate is simply due to chance. In that case the effect is said to be not statistically significant.

References

Carlin, J., Taylor, P. and Nolan, T. (1998). 'School based bicycle safety education and bicycle injuries in children: a case-control study', *Injury Prevention*, **4**, 1, 22–27 [online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1730310/pdf/v004p00022.pdf [3 December, 2014]..

Colwell, J. and Culverwell, A. (2001). 'An examination of the relationship between cycle training, cycle accidents, attitudes and cycling behaviour among children', *Ergonomics*, **45,** 9, 640–648.

Department for Transport (2013). *Reported Road Casualties in Great Britain: Main Results 2012* (Statistical Release 27 June). London: Department for Transport [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/245754/reported-road-casualties-in-great-britain-main-results-2012.pdf [3 December, 2014].

Ducheyne, F., De Bourdeaudhuij, I., Lenior, M. and Cardon, G. (2013). 'Does a cycle training course improve cycling skills in children?', *Accident Analysis and Prevention*, **59**, 38–45.

Education Endowment Foundation (2014). *Evaluation Glossary* [online]. Available: http://educationendowmentfoundation.org.uk/evaluation/evaluation-glossary/ [3 December, 2014].

Heesch, K., Garrard, J. and Sahlqvist, S. (2011). 'Incidence, severity and correlates of bicycling injuries in a sample of cyclists in Queensland, Australia', *Accident Analysis and Prevention*, **43**, 2085–2092.

Ipsos MORI (2010). *Research to Explore Perceptions and Experiences of Bikeability Training Amongst Parents and Children* [online]. Available: http://bikeability.dft.gov.uk/wp-content/uploads/perceptions-and-experiences-of-bikeability.pdf [3 December, 2014].

Johnson, M., Newstead, S., Charlton, J. and Oxley, J. (2011). 'Riding through red lights: the rate, characteristics and risk factors of non-compliant urban commuter cyclists', *Accident Analysis and Prevention*, **43**, 323–328.

Klin, B., Rosenfeld-Yehoshua, N., Abu-Kishk, I., Efrati, Y., Kozer, E., Jeroukhimov, I., Eshel, G. and Lotan, G. (2009). 'Bicycle-related injuries in children: disturbing profile of a growing problem', *Injury: International Journal of the Care of the Injured*, **40**, 1011–1013.

Knowles, J., Adams, S., Cuerden, R., Savill, T., Reid, S. and Tight, M. (2009a). *Collisions Involving Cyclists on Britain's Roads: Establishing the Causes* (Report No. PPR 445). Bracknell: IHS [online]. Available: http://www.worthingrevolutions.org.uk/sites/worthingrevolutions.org.uk/files/PPR445.pdf [3 December, 2014].

Knowles, K., Adams, S., Cuerden, R., Savill, T., Reid, S. and Tight, M. (2009b). *Collisions Involving Pedal Cyclists on Britain's Roads: Establishing the Causes. Technical Annex to PPR 445*. Bracknell: IHS [online]. Available: http://www.trl.co.uk/online_store/reports_publications/trl_reports/cat_road_user_safety/report_technical_annex_to_ppr445___collisions_involving_pedal_cyclists_on_britains_roads:_establishing_the_causes.htm [3 December, 2014]

Maring, W. and van Schagen, I. (1990). 'Age dependence of attitudes and knowledge in cyclists', *Accident Analysis and Prevention*, **22**, 2, 127–136.

McLaughlin, K. and Glang, A. (2010). 'The effectiveness of a bicycle safety program for improving safety-related knowledge and behavior in young elementary students', *Journal of Pediatric Psychology*, **35**, 4, 343–353.

Miller, P. (2012). The use of conspicuity aids by cyclists and the risk of crashes involving other road users: a population based case-control study. Unpublished Doctorate Thesis, University of Nottingham [online]. Available: http://etheses.nottingham.ac.uk/2855/1/NBAS_Thesis_FOR_RESUBMISSION_Oct_2012_FINAL_PRINT.pdf [3 December, 2014].

OECD/International Transport Forum (2013). *Cycling, Health and Safety*. Paris: OECD Publishing.

Plumert, J.M., Kearney, J.K., Cremer, J F., Recker, K.M. and Strutt, J. (2011). 'Changes in children's perception-action tuning over short time scales: bicycling across traffic-filled intersections in a virtual environment', *Journal of Experimental Child Psychology*, **108**, 2, 322–337.

Plumert, J.M., Kearney, J.K. and Cremer, J.F. (2004). 'Children's perception of gap affordances: bicycling across traffic-filled intersections in an immersive virtual environment', *Child Development*, **75**, 4, 1243–1253.

Royal Society for the Prevention of Accidents (2001). *The Effectiveness of Cyclist Training* [online]. Available: http://www.rospa.com/roadsafety/info/cyclist_training_effectiveness.pdf [3 December, 2014].

Savill, T., Bryan-Brown, K. and Harland, G. (1996). *The Effectiveness of Child Cycle Training Schemes*. Crowthorne: Transport Research Laboratory.

Simpson, A.H. and Mineiro, J. (1992). 'Prevention of bicycle accidents', *Injury*, **23**, 3, 171–173.Cited in: Ducheyne, F., De Bourdeaudhuij, I., Lenoir, M. and Cardon, G. (2013). 'Does a cycle training course improve cycling skills in children? *Accident Analysis and Prevention*, **59**, 38– 45.

Turner, S., Binder, S. and Roozenburg, A. (2009). *Cycle Safety: Reducing the Crash Risk* (Research Report No. 389). Wellington: NZ Transport Agency [online]. Available: http://www.nzta.govt.nz/resources/research/reports/389/ [3 December, 2014].

van Schagen, I. and Brookhuis, K. (1994). 'Training young cyclists to cope with dynamic traffic situations', *Accident Analysis and Prevention*, **26**, 2, 223–230.

Further reading

Ayres, T. (2010). 'HF/E professionals: help make biking to school safer', *Ergonomics in Design*, Summer, 30–31.

Bikeability (n.d.) *Bikeability Delivery Guide: Delivery Guidance for Instructors and Training Providers*. London: Department for Transport [online]. Available: http://bikeability.dft.gov.uk/wp-content/uploads/Bikeability_Delivery_Guide.pdf [4 December, 2014].

Bikeability (2014). *What is Bikeability?* [online]. Available: https://bikeability.dft.gov.uk/what-is-bikeability/ [5 December, 2014].

Briem, V., Radeborg, K., Salo, I. and Bengtsson, H. (2004). 'Developmental aspects of children's behavior and safety while cycling', *Journal of Pediatric Psychology*, **29**, 5, 369–377.

British Cycling (n.d.). *Road Safety.* Manchester: British Cycling.

Chihak, B. J., Grechkin, T. Y., Kearney, J. K., Cremer, J. F., & Plumert, J. M. (2014). How children and adults learn to intercept moving gaps. *Journal of experimental child psychology*, *122*, 134-152.

CTC (n.d.). *Safety in Numbers in England* [online]. Available: http://www.ctc.org.uk/sites/default/files/0905_sin_full_rpt_0.pdf [3 December, 2014].

CTC (2012). *Cycle Digest*, **69** (whole issue) [online]. Available: http://www.ctc.org.uk/sites/default/files/file_public/cycledigest692012.pdf [3 December, 2014].

CTC (2012). *Response to the Inquiry into the Strategic Framework for Road Safety* [online]. Available: http://www.ctc.org.uk/sites/default/files/1110_ctc_transcom_rss-con-final.pdf [3 December, 2014].

CTC (2013). *Submission to the Commons Transport Committee for its Evidence Session on Cycle Safety Further to the Committee's 2012 Inquiry on the Government's Road Safety Strategy* [online]. Available: http://www.ctc.org.uk/sites/default/files/1311_rg_transcom_cycle-safety_con.pdf [3 December, 2014].

CTC (2013). *Cycle Safety* [online]. Available: http://www.ctc.org.uk/sites/default/files/1311_rg_m-macleod-cycle-safety-debate_brf.pdf [3 December, 2014].

Department for Transport (2008). *Cycle Infrastructure Design* (Local Transport Note 2/08). London: TSO [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/3808/ltn-2-08.pdf [3 December, 2014].

Department for Transport (2009). *Collisions Involving Cyclists on Britain's Roads: Establishing the Causes. Findings* (Report PPR 445). London: Department for Transport [online]. Available: http://webarchive.nationalarchives.gov.uk/20090417002224/http:/www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme1/ppr445.pdf [3 December, 2014].

Department for Transport (2010). *Cycling, Safety and Sharing the Road: Quantitative Research with Cyclists and Other Road Users* (Road Safety Web Publication No. 17). London: TSO [online]. Available: http://webarchive.nationalarchives.gov.uk/20111005175849/http://assets.dft.gov.uk/publications/safety-cycling-and-sharing-the-road-qualitative-research-with-cyclists-and-other-road-users/rswp17.pdf [3 December, 2014].

Department for Transport (2011). *Strategic Framework for Road Safety*. London: Department for Transport [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/8146/strategicframework.pdf [3 December, 2014].

Department for Transport (2011). *Infrastructure and Cyclist Safety* (TRL Report PPR 580). London: TSO [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/3731/infrastructure-and-cyclist-safety.pdf [3 December, 2014].

Department for Transport (2012). *Reported Road Casualties in Great Britain: 2011 Annual Report. Hospital Admissions Data on Road Casualties in England*. London: Department for Transport [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/9279/rrcgb2011-06.pdf [3 December, 2014].

Department for Transport (2013). *Report Road Casualties in Great Britain: Quarterly Provisional Estimates Q2 2013*. London: Department for Transport [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/277527/road-accidents-and-safety-quarterly-estimates-q3-2013.pdf [3 December, 2014].

Department for Transport (2013). *Reported Road Casualties in Great Britain: Guide to the Statistics and Data Sources*. London: Department for Transport [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/259012/rrcgb-quality-statement.pdf [3 December, 2014].

Dragutinovic, N. and Twisk, D. (2006). *The Effectiveness of Road Safety Education*. Leidschendam, NL: SWOV Institute for Road Safety Research [online]. Available: http://www.swov.nl/rapport/r-2006-06.pdf [3 December, 2014].

Ducheyne, F., De Bourdeaudhuij, I., Lenoir, M., Spittaels, H. and Cardon, G. (2013). 'Children's cycling skills: development of a test and determination of individual and environmental correlates', *Accident Analysis and Prevention*, **50**, 688–697.

European Transport Safety Council (2011). *Unprotected Road Users – a Key Concern of Road Safety* (Road Safety Performance Index Flash 19). Brussels: European Transport Safety Council [online]. Available: http://www.etsc.eu/documents/ETSC_Flash19_unprotected_road_users.pdf [3 December, 2014].

Education Endowment Foundation and The Sutton Trust (2013). *Sutton Trust - EEF Teaching and Learning Toolkit* [online]. Available: http://educationendowmentfoundation.org.uk/uploads/toolkit/Teaching_and_Learning_Toolkit_(Spring_2013).pdf [4 December, 2014].

Everett, P. (2013). *The City of London's Road Danger Reduction Plan: Supporting Technical Document*. London: City of London Corporation [online]. Available: https://www.cityoflondon.gov.uk/services/transport-and-streets/traffic-management/road-danger-reduction/Documents/road-danger-reduction-plan-2013-supporting-technical-document.pdf [3 December, 2014].

Goldacre, B. and Spiegelhalter, D. (2013). 'Bicycle helmets and the law: Canadian legislation had minimal effect on serious head injuries', *BMJ*, **346.**

Great Britain. Parliament. House of Commons. Transport Committee (2012). *Road Safety. Report, Together with Formal Minutes, Oral and Written Evidence* (HC 506). London: TSO [online]. Available: http://www.publications.parliament.uk/pa/cm201213/cmselect/cmtran/506/506.pdf [3 December, 2014].

Great Britain. Parliament. House of Commons. Transport Committee (2012). *Road Safety: Government Response to the Committee's Second Report of Session 2012–13* (HC 648). London: TSO [online]. Available: http://www.publications.parliament.uk/pa/cm201213/cmselect/cmtran/648/648.pdf [3 December, 2014].

Grechkin, T.Y., Chihak, B.J., Cremer, J F., Kearney, J.K. and Plumert, J.M. (2013). 'Perceiving and acting on complex affordances: How children and adults bicycle across two lanes of opposing traffic', *Journal of Experimental Psychology: Human Perception and Performance*, **39**,1, 23.

Hatfield, J. (n.d.). *Review of Evaluation of Education as a Countermeasure for Cyclist Injury* [online]. Available: http://www.irmrc.unsw.edu.au/documents/JHatfield.pdf [3 December, 2014].

Hooshmand, J., Hotz, G., Neilson, V. and Chandler, L. (2014). 'BikeSafe: evaluating a bicycle safety program for middle school aged children', *Accident Analysis and Prevention*, **66**, 182–186.

Keep, M. (2013). *Road Cycling: Statistics* (House of Commons Library Standard Note SN06224) [online]. Available: http://www.parliament.uk/briefing-papers/SN06224/road-cycling-statistics [19 February, 2014].

Krag, T. (2005). 'Cycling, safety and health.' In: Allsop, R., Beckmann, J. and Mackay, G. (Eds). *ETSC Yearbook 2005: Safety and Sustainability*. Brussels: European Transport Safety Council [online]. Available: http://www.etsc.eu/documents/Yearbook_2005.pdf [3 December, 2014].

London Assembly Transport Committee (2012). *Written Submissions Received for the Transport Committee's Investigation into Cycle Safety in London*. London: London Assembly Transport Committee [online]. Available: http://www.cycling-embassy.org.uk/sites/cycling-embassy.org.uk/files/documents/Cycle%20Safety%20Evidence%20reduced.pdf [3 December, 2014].

London Assembly Transport Committee (2012). *Gearing Up: an Investigation into Safer Cycling in London*. London: Greater London Authority [online]. Available: http://www.london.gov.uk/sites/default/files/For%20print%20-%20Gearing%20up%20-%20safer%20cycling%20in%20London%2029%20Nov.pdf [3 December, 2014].

Macarthur, C., Parkin, P., Sidky, M. and Wallace, W. (1998). 'Evaluation of a bicycle skills training program for young children: a randomized controlled trial', *Injury Prevention*, **4**, 2, 116–121.

Margolis, S. (2011). 'Free cycle training for adults – does it make a difference?' *Cycle Digest*, **68,** 4 [online]. Available: http://www.ctc.org.uk/sites/default/files/file_public/cycledigest682011-12.pdf [3 December, 2014].

Mehan, T., Gardner, R., Smith, G. and McKenzie, L. (2009). 'Bicycle-related injuries among children and adolescents in the United States', *Clinical Pediatrics*, **48**, 2, 166–173.

National Audit Office (2009). *Improving Road Safety for Pedestrians and Cyclists in Great Britain*. London: TSO [online]. Available: http://www.nao.org.uk/wp-content/uploads/2009/05/0809437.pdf [3 December, 2014].

OECD/International Transport Forum (2012). *Cycling Safety: Key Messages.* Paris: OECD Publishing.

Richmond, S.A., Yu, J.Z., Stoer, A., Howard, A. and Macarthur, C. (2014). 'Prevention of bicycle-related injuries in children and youth: a systematic review of bicycle skills training interventions', *Injury Prevention*, **20**, 3,191–195.

Sustrans (n.d.). *20mph: a Snapshot of the Evidence*. Bristol: Sustrans.

Stevens, E., Plumert, J. M., Cremer, J.F. and Kearney, J.K. (2013). 'Preadolescent temperament and risky behavior: bicycling across traffic-filled intersections in a virtual environment', *Journal of Pediatric Psychology*, **38**, 3, 285–295

Transport for London (2011). *Pedal Cyclist Collisions and Casualties in Greater London* (Topic Factsheet 2011-1). London: Transport for London [online]. Available: http://www.tfl.gov.uk/assets/downloads/pedal-cyclist-collisions-and-casualities-in-greater-london-sep-2011.pdf [3 December, 2014].

Transport for London (2013). *Collisions and Casualties on London's Roads: Annual Report 2012*. London: Transport for London [online]. Available: http://www.tfl.gov.uk/cdn/static/cms/documents/collisions-and-casualties-2012.pdf [3 December, 2014].

Turner, S., Roozenburg, A. and Smith, A. (2009). *Roundabout Crash Prediction Models* (Research Report No. 386). Wellington: NZ Transport Agency [online]. Available: http://www.nzta.govt.nz/resources/research/reports/386/docs/386.pdf [3 December, 2014].

VicRoads (n.d.). *Pedestrian and Cyclist Safety in Victoria*. Kew, Vic: VicRoads.

NFER provides evidence for excellence through its independence and insights, the breadth of its work, its connections, and a focus on outcomes.

- independent
- insights
- breadth
- connections
- outcomes

NFER ref. BIKE

ISBN. 978-1-910008-46-1

National Foundation for Educational Research
The Mere, Upton Park,
Slough, Berks SL1 2DQ

T: 01753 574123
F: 01753 691632
E: enquiries@nfer.ac.uk
www.nfer.ac.uk