TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

MASTERARBEIT

An Empirical Evaluation of 2D and Interactive 3D Terrain Visualisations for Cycling Maps

Ausgeführt am Department für
Geodäsie und Geoinformation
der Technischen Universität Wien

unter der Anleitung von
Univ.Prof. Mag.rer.nat. Dr.rer.nat. Georg Gartner, TU Wien

und
Dipl.-Ing. Florian Ledermann, TU Wien

durch

Myles Edward Cook

Burggasse 88/22, 1070 Wien

15. Mai 2017

Unterschrift (Student)

MASTER'S THESIS

An Empirical Evaluation of 2D and Interactive 3D Terrain Visualisations for Cycling Maps

Conducted at The Department of
Geodesy and Geoinformation
Technical University Vienna

under the supervision of
Univ.Prof. Mag.rer.nat. Dr.rer.nat. Georg Gartner, TU Wien

and
Dipl.-Ing. Florian Ledermann, TU Wien

by

Myles Edward Cook

Burggasse 88/22, 1070 Wien

15th May 2017

Signature (Student)

Acknowledgments

Enormous, boundless thanks must firstly go to my thesis supervisor, Dipl.-Ing. Florian Ledermann, for his endless support, patience, guidance, good nature, and dissertation-saving technical expertise. Likewise to Prof. Dr. Georg Gartner, for his indispensible feedback and input throughout the course of my thesis. I am indebted to Silvia Klettner for statistical advice, and indeed to all other members of the Cartography Research Group at the Technical University of Vienna – for acting as participants, for providing feedback, and for being warm and friendly whenever I was working in the corner of the office.

To everybody involved in the International Cartography Masters Programme, and the DAAD, who provided me with financial support – I must say thankyou for giving me the opportunity to be a student on a unique and wonderful master's programme. My further gratitude goes to Juliane Cron, coordinator of the Cartography MSc, for her great support throughout the masters.

Finally, to my family back in the UK, and the irreplaceable, inimitable group of friends I have across Europe, I say the biggest thank you of all. I couldn't have done it were it not for your tremendous support – and I promise that now I'll talk about something other than my thesis!

Abstract

The last two decades have borne witness to considerable growth in the availability of interactive and 3D cartographies. However, questions about their true usefulness and usability remain, with the results of associated research drawing conflicting conclusions. This knowledge deficit is especially apparent within the design of cycling maps and cycle route-planners. 3D cartography has enabled new forms of terrain visualisation for these use scenarios, but their usability relative to traditional methods of terrain visualisation is not yet fully understood. In response to this deficit, this thesis aimed to assess the relative usability (measured in terms of efficiency, effectiveness and user-preference) of 2D and 3D elevation visualisations for cycle route-planners.

In order to fulfil this aim, an empirical user-study was conducted with 36 participants. Participants were asked to solve a range of typical cycle route-planning tasks (related to height detection, slope detection and climb estimation), for a variety of 2D (arrow, colour, elevation profile) and interactive 3D (3D elevation profile and 3D terrain model) elevation visualisations. Study participants also provided feedback on their visualisation preference for each of the tasks. These usability factors were assessed using a digital survey, which allowed inclusion of interactive 3D visualisations within a controlled experimental setting, whilst also automatically logging user response times, user answers, and preference feedback.

The findings demonstrated significant usability differences between individual 2D and 3D elevation visualisations, although it was not possible to broadly state that one dimensionality was more or less 'usable' than the other. Further, the usability of each visualisation was found to be strongly dependent upon the type of route-planning task. However, the 2D elevation profile was most efficient in the widest range of scenarios, and those (generally 3D) visualisations which demanded interaction or high levels of cognitive processing were least efficient. The results also showed that those visualisations which placed the lowest cognitive load on the user were most effective; this factor appears to have had a greater influence on efficacy than the dimensionality of the visualisations. For all task types, users preferred 2D visualisations, most especially the 2D elevation profile.

Contents

List of Figures

List of Tables

1 Introduction

For almost as long as there have been bicycles, there have been cycling maps. It was in the latter part of the 19th century that the invention of the safety bicycle, with its low riding position and chain driven rear-wheel, led to an explosion in the popularity of cycling (Herlihy 2004; Rubinstein 1977). The new and rapidly expanding sector of society that resulted – 'the cyclist' – naturally also demanded a new breed of cartography. In Europe and North America, there followed the development and proliferation of maps which communicated the information most pertinent to the early generation of cyclists, such as navigable routes, the quality of road surfaces and the steepness of hills (Nicholson 2004; Akerman 2002).

Viewed in the construct of the map communication model (Robinson & Petchenik 1975), such maps may be framed within a 'traditional' cartographic structure: they present and communicate a selection of spatial information – information desired by a cyclist – by means of a visualisation. The cartographer producing a cycling map therefore has two goals: determining which information is necessary for a cyclist; and determining how best to convey that information visually. For the cyclist, this information is most frequently used to support decision making in route planning, navigation and wayfinding (Brügger 2015; Ehlers et al. 2002; Hochmair 2004).

This model of map communication has fallen out of favour within the last two decades, concurrent partly with a growth in the appreciation of the social context of cartographic interpretation, and also following the increasing interactivity of cartographic products (e.g. Crampton & Krygier 2006; MacEachren 1994; MacEachren & Kraak 2011). Nonetheless, the likes of Adams (2009) and Brodersen (2001) maintain that the broad paradigm of maps existing as a means of information communication remains fundamental to the discipline of cartography. Crucially, this approach also continues to underpin functional research focussing on cartographic design (Montello 2002; Kitchin et al. 2009).

How can this model of information transfer aid the design of contemporary cycling maps and cycle route-planners? The choice of information encoded within a cycling map is inextricably linked to its target audience and specific purpose (Wessel & Widener 2015; Sherwin & Bartle 2012). Likewise, the most important criteria in determining route choice vary significantly between different cyclists (Hochmair 2004; Sherwin & Bartle 2012). For example, whilst some cyclists may be most interested in the location of dedicated cycle paths, others may wish primarily to avoid areas of congested or fast traffic. However, multiple authors have suggested that terrain features, in particular route steepness and elevation gain, are crucial factors for the vast majority of cyclists (e.g. Brügger 2015; Hochmair 2005; Winters et al. 2011). From a cartographic design perspective, the question therefore arises: what are the optimum visual means for communicating this terrain information?

Brügger (2015) began to address this question, with an empirical study into the efficacy of elevation depictions in 2D city cycling-maps. However, the issue is far from resolved. In particular, the growth of interactive, mobile, and 3D cartography in recent years has presented a myriad of new cartographic opportunities for communicating elevation information to the cyclist (e.g. Bleisch & Dykes 2008; Incoul et al. 2015; Liu 2015). But do these new methods offer any advantages over existing static 2D depictions? Are they more usable for the cyclist – that is to say, are they more efficient and effective at communicating terrain information, and more preferable for the user? This research project is a response to those questions.

1.1 *Justification*

In seeking answers to clarify these issues, this thesis aims in part to directly answer the calls of Brügger et al. (2016) for further empirical research into the usability of interactive cycling maps. However, although the focus of this study lies explicitly on cycle route-planners, the need for further research into the general usability and desirability of 3D spatial visualisations is apparent throughout wider cartography. Though there exists a broad assumption that such visualisations *should* be more intuitive and user-friendly than their 2D counterparts (e.g. Bleisch & Dykes 2008; Popelka & Brychtova 2013), research aimed at testing this assumption is both limited in its scope, and conflicting in its conclusions (Section 3.5). Moreover, Herman & Stachon (2016) highlighted that of the few empirical studies which do compare 2D and 3D spatial visualisations, the majority were conducted using static rather than interactive depictions. The results of these studies may therefore not apply to the interactive visualisations users are increasingly likely to encounter (Herman & Stachon 2016). Likewise, the work of Liao et al. (2016) and Savage et al. (2004), amongst many others, suggests that the relative benefits of 2D and 3D spatial visualisations are heavily dependent on use context and task. As such, further research looking specifically at the usability of interactive 3D visualisations from a cycle route-planning perspective is desirable.

Although this thesis partly aims to fill a research gap within the realm of 3D cartography, it also aims to add to the small body of work surrounding the cartographic design of cycling maps in general. As a number of authors have noted, there exists a relative dearth of academic literature focussing on this topic (Brügger 2015; Dickinson 2012; Wessel & Widener 2015), the result of which has been difficulty in designing effective cycling maps (Sherwin & Melia 2012; Wessel & Widener 2015). By considering the usability of 3D visualisations in a cycling context, the research outcomes of this study may aid in the design of cartographic products which are more user-oriented, and thus an embodiment of what Norman (2002) termed 'user-centred design'. The possibility thereafter is to produce a product that more fully meets a user's wants and needs (Nivala 2005). Designing more user-friendly cycling maps and route planners is an especially important task, given that many urban governments are attempting to encourage cycling uptake as a means to fight urban congestion, greenhouse-gas emissions, and even obesity (Maibach et al. 2009; Sherwin & Melia 2012; Su et al. 2010). Well-designed cycling maps, which effectively communicate information such as elevation to the user, have been purported to encourage and enable cycling (Hochmair 2005; Sherwin & Bartle 2012). 3D visualisations have the potential to provide this information more efficiently, effectively, and satisfyingly than 2D maps – and in doing so encourage cycling even more. However, further research is needed to assess whether or not that is indeed the case.

1.2 *Broad Research Aim*

In light of the above, the broad aim of this thesis is therefore to assess the relative usability of 2D and 3D elevation visualisations for cycle route-planners. Usability in this context is defined according to the ISO Standard 9241-11, in terms of efficiency, effectiveness and satisfaction (Bevan et al. 2015). In striving for this aim, the study further builds on the work of Brügger (2015) and Brügger et al (2016), to add to the small existing pool of empirical research into the cartographic design of cycling maps and cycle route-planners.

The specific research questions which support this broad aim are strongly influenced by existing literature on the subject. Hence, they are included after the proceeding literature review.

2 Thesis Structure

The structure of the subsequent thesis aims to logically present the design, execution and analysis of a cartographic user-study, which was conducted to address the aim of the thesis. In Chapter 3, the contextual and scientific background of the study is introduced, including a brief look at the history of related map-design research. This section also latterly includes a review of map-design research approaches, in addition to a review of existing cycle-map elevation depictions, found both within academia and in published media (print and digital). In the context of this literature, Chapter 4 then presents the primary research questions which run throughout the thesis. In Chapter 5, the design of the empirical study used to assess the usability of 2D and 3D elevation visualisations is presented and discussed. The results of this user study are then presented, without commentary, in Chapter 6. Chapter 7 discusses these results in greater depth, and in relation to the wider scientific and societal context of the research (as was introduced in the earlier literature review). Limitations and potential flaws of the study are covered in Chapter 8, before some final conclusions are presented in Chapter 9.

3 Scientific Background and Literature Review

Broadly speaking, this study falls into the cartographic sub-discipline of cognitive map-design – a field defined by Montello (2002) as research with the goal of "understanding human cognition in order to improve the design and use of maps." Subscribing to this definition, the essence of this work may thusly be considered as a study of the human cognition of 2D and 3D elevation visualisations, with the goal of improving the design and usability of cycling maps and cycle route-planners. Montello (2002) provided a detailed history of cognitive map-design research. However, a brief overview is provided below (Section 3.1), followed by a review of associated research approaches (Section 3.2), in order to place the thesis in a wider scientific context.

Thereafter, the focus of the literature review narrows to consider the small body of empirical research which surrounds cycling maps specifically (Section 3.3). Given the paucity of academic literature on this topic, this section is supplemented by a subsequent review of elevation depictions in existing cycling maps and cycle route-planners (Section 3.4). Latterly, the emphasis narrows yet further, to critique previous research into the relative usability of 2D and 3D cartographic depictions (Section 3.5). Whilst the initial sections of the review aim to place the study in a broad historical and theoretical context, the latter are intended to provide the detailed background necessary to support the thesis research questions posed in Chapter 4, and the methodology developed in Chapter 5.

3.1 *Cognitive Map-Design*

Given the long and complex history of cartography, academic research aimed at understanding human cognition of maps, and thus at improving map design was – at least until the middle of the twentieth century – surprisingly lacking. Until this time, map-design was arguably a pursuit guided mainly by intuition, practicality, and incremental improvements in knowledge (Ciolkosz-Styk 2012; Zyszkowska 2015).

However, that is not to say that all academic cartographers prior to the 1950s totally overlooked the ways in which users perceived maps – and by association the ways in which the design of maps could be improved to aid the perception of those users. Two threads of research from the German-speaking world support this statement. Around 1898, the Viennese cartographer Karl Peucker was amongst the first to consider the physiology of the human eye, when developing hypsometric colour schemes for relief depiction (Speich 2009; Patterson & Jenny 2011). Likewise, the seminal work of Eckert (1921-25), '*Die Kartenwissenschaft*' ('Map-Science'), not only helped herald the new era of 'scientific cartography', but also stressed that successful map design must take into account the psychological perceptions of the user (Scharfe 1986; Montello 2002). Indeed, Eckert (1925:551), cited in Scharfe (1986), suggests that a cartographer and psychologist working together would result in "extraordinary progress" in map design.

Nonetheless, it was only with the publication of the *The Look of Maps* (Robinson 1952), that cognitive map-design research truly rose to prominence, and latterly flourished into a widely accepted academic pursuit (Montello 2002; Zyszkowska 2015; Crampton & Krygier 2005). In *The Look of Maps*, Robinson (1952) suggested that map design, and the needs of the user, should be awarded as much importance as the data behind the map. In some senses, Robinson's work draws on the 'scientific' calls of Eckert (1925), by further suggesting that good map design must be rooted in empirical research, to determine how different map design choices impact the communication of information, via the psychological perceptions of the map reader (Olson 2005; Montello 2002). These design decisions, Robinson suggests, would be best informed by systematic, functional and objective studies into the human cognition of maps. The development of the map communication model (Figure 3-1), whereby maps are characterised as a means of spatial information transfer, actively encouraged such empirical studies (Edney 2005). By defining cartography as a form of communication, the model helped legitimise empirical studies that – ultimately – assessed the effectiveness of communication (Edney 2005; Zyszkowska 2015). Thus, *The Look of*

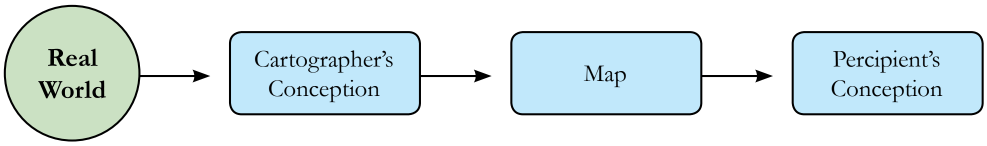


Figure 3-1 *A simplified version of the map communication model, demonstrating the transfer of information from the 'real world' to the map reader (the 'percipient'), through the filters of the cartographer's conception and the map itself. Based on Robinson and Petchenik (1975)*

Maps helped to catalyse a new wave of empirical research and dedicated scholarly pursuit into effective map design, which would remain prominent in Cartography into the 1980s (Edney 2005; Montello 2002).

Many of the studies which initially followed The Look of Maps were devoted to analysis of the perception of map symbols, especially for thematic maps (Montello 2002; Zyszkowska 2015; MacEachren 1994). These studies often took a 'psychophysical approach', whereby the human response (psycho–) to a changing symbol (–physical) was studied (Ciolkosz-Styk 2012). A large body of research thusly assessed, and attempted to quantify, the difference between human perception of the value of graduated symbols, and the actual measurable size of those symbols (Montello 2002). Often cited in this respect is the work of (Flannery 1956), who found that the area of a symbol perceived by a map user is a power function of the actual area. Similar work attempted study the perception of pseudo-3D cartographic symbols (Ekman et al. 1963), while other research found that perceptual differences were not apparent for certain symbols, such as squares (Crawford 1973). Beyond graduated symbols, later work went on to study the human perception of many other 'small-scale' cartographic design features, including grey-tones, typefaces and colour hues (Ciolkosz-Styl, 2012).

A primary justification behind such studies, was that if psychophysical relations and perceptual anomalies could be identified, they could also be accounted for in map design. The resultant map should therefore more effectively and efficiently communicate its information to the map reader – without perceptual discrepancies or confusion – in much the same way

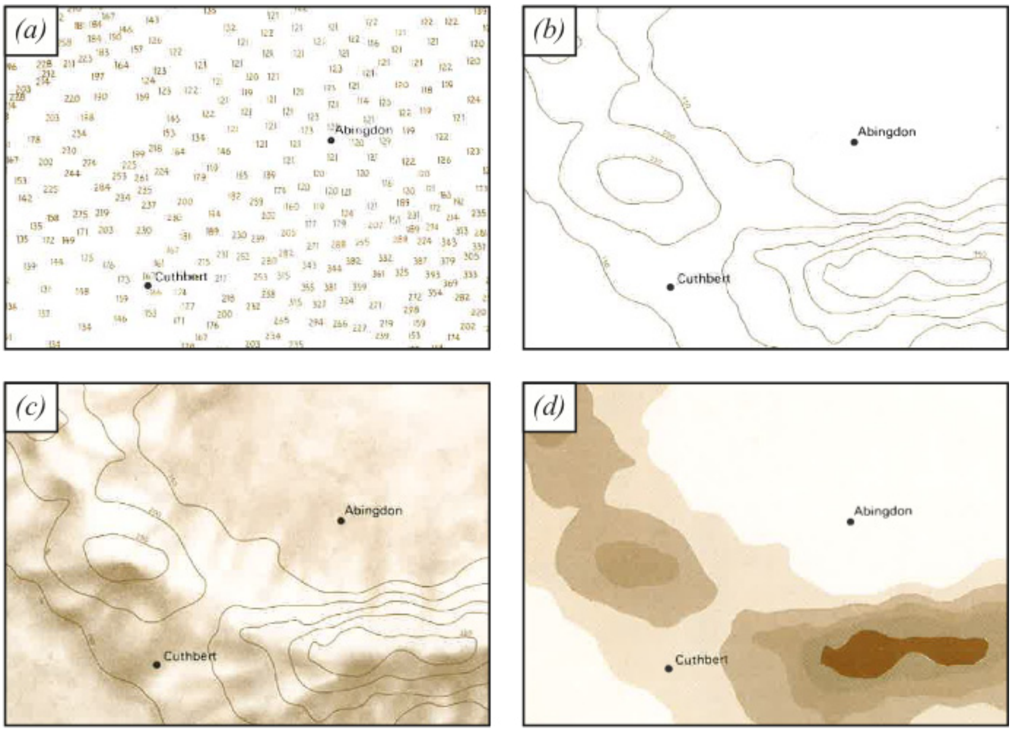


Figure 3-2 *Sample portions of the test stimuli used in Phillips et al. (1975), showing: (a) spot height, (b) contour, (c) contour with hill shading, and (d) hypsometric shading elevation depictions. Adapted from Phillips et al. (1975).*

that this thesis is attempting to determine the most effective, efficient and desirable visual means for communicating height information to a cyclist.

However, more directly relevant to the aim of this thesis is the smaller body of early empirical research focussing on topographic maps, which was concerned especially with cartographic depictions of the third dimension, i.e. relief (Phillips 1979). Early research was conducted by Phillips et al. (1975), who attempted to determine the 'legibility' of different forms of relief depictions – namely contours, hill-shaded contours, hypsometric tints and spot-heights (Figure 3-2). The researchers tested 175 participants on a series of questions aimed at imitating real-life map use, such as finding the highest and lowest points on a map, and determining the shortest route that remained above a certain height. The legibility of each map type was measured in terms of the speed and accuracy of participant answers (Phillips et al., 1975). Though the study found significant differences between the legibility of map types, results were highly use-dependent: no single studied map type was universally more 'legible' than another. This reinforces the need for map-design research that focuses on design issues for particular use situations, such as cycle route-planning. Subsequent work applied similar methodologies to early computer generated maps, including early perspective-3D depictions (Phillips & Noyes 1978), and colour schemes for hypsometric maps (Phillips 1982).

Throughout the 1970s and 80s, further research in this vain continued to gradually provide small cartographic design recommendations (e.g. Castner & Wheate 1979; Griffin & Lock 1979; Shurtleff & Geiselman 1986) though, as Petchenik (1983) points out, practical implementations of these recommendations are harder to find. Indeed, some have suggested that even where clear conclusions and map-design guidelines did result from empirical research, their use was limited (Petchenik 1983; Montello 2002). However, Slocum et al. (2005), cited in Ciolkosz-Styl (2012), point out that a lack of adoption alone should not damage the status of map-design research, or the insight that it produces.

Nonetheless, closer scrutiny has also showed many of these early studies to suffer from methodological flaws, which raise doubts about the validity of their conclusions. For example, in Phillips et al. (1975), the ordering of map-stimuli may have promoted learning effects that influenced study outcomes. Indeed, poorly designed methodologies may explain why similar empirical map-design research sometimes led to wildly different results (Petchenik, 1983). For example, in studying the human perception of graduated circles, the likes of Flannery (1956), Ekman & Junge (1961) and Crawford (1973), each determined notably different results for the degree to which the size of the circles was over or under-estimated. These differing results emphasise the importance of developing rigorous methodologies when conducting empirical map-design research.

A further criticism is that the maps in empirical studies were often unrepresentative of real-life use situations, having been excessively simplified in order to isolate certain design variables (Petchenik, 1983). These maps have been described as ‘quasi-maps’, and are devoid of the level of cartographic context and detail found in ‘real-maps’ (Ciolkosz-Styl, 2012). Certainly, this is an issue evident in the research of Phillips (1979) and Griffin and Lock (1979), amongst others. Likewise, the tasks by which usability was assessed were often similarly simplistic and unrepresentative of real life scenarios (Ciolkosz-Styl, 2012). Though such simplification is indeed necessary in order to gain insight into specific design features, it reduces the validity and applicability of research outcomes. Subsequent conclusions are arguably context-dependent, simplistic, and unrepresentative, with widely applicable design recommendations perhaps an unattainable goal (Worth 1989; Zyszkowska 2015).

Similarly, Montello (2002) notes that individual map-users will always interact with and experience maps in an individual way. As such, striving for general design conclusions is perhaps a futile goal when conducting map-design research. However, a similar criticism could be levelled against any piece of user-centred design research. The importance is therefore to ensure that the participants being studied are sufficiently representative of the end-users of the map in question. Rather than invalidating all cognitive map-design research, these issues instead primarily highlight the necessity of developing a rigorous and scientifically valid methodology.

Beyond methodological problems, the perceived value of this form of empirical research fell further during the 1980s, concurrent with increasing criticism levelled toward the map communication model. The model was criticised as failing to appreciate the knowledge, diversity and context of map users, alongside the increasing interactivity of map communication (Crampton & Krygier 2005; MacEachren 1994). By this point, maps were increasingly considered 'social constructions' (Crampton 2001) – a definition which rendered the oft socially-unconsidered and arguably simplistic conclusions of laboratory research inapplicable to real-life map use. If maps were no longer seen as single-purpose forms of linear communication, then 'map-engineering'-esque empirical research, which studied them as such, also appears invalid (MacEachren 1994). Montello (2002) further notes a general mood amongst cartographers that empirical studies also often produced conclusions 'that we already knew', and did little to add to cartographic design knowledge as a whole.

Even where design recommendations based on empirical research have been incorporated into mainstream cartography, they have not been immune to criticism. For example, the perceptually adjusted graduated circles proposed by Flannery (1956), which remain a feature of ESRI's ArcGIS software today, have been criticised for their lack of measurable accuracy by esteemed eartographer Edward Tufte (Krygier 2007), and for their user-specific nature by others (Griffin 1985).

In spite of these criticisms, in recent years many authors have talked of a resurgence in cognitive map-design research (Montello, 2002; Zyskowska, 2015; Ciolkosz-Styl, 2012). Similarly, others have talked of the vital necessity of cognitive map-design in modern cartographic research (Kitchin et al. 2009; MacEachren 2013).

In part, this stems from the increasing importance placed upon user-centred design – with map-design research aimed at improving the usability of cartographic products consequently gaining prominence (Nivala 2005). It also results from the development of new cartographic technologies: digital, 3D, interactive, animated, and mobile cartography have all come to the fore in recent decades (Incoul et al. 2015). As these technologies do not enjoy the same long history as traditional cartographic forms, there is much opportunity – and much need – for empirical research to tell us 'something we don't already know' (Montello 2002; MacEachren 1994; Shepherd 2007; Çöltekin et al. 2009; Roth 2013).

Beyond being the source of a new cartographic 'knowledge gap', these new technologies also provide some solutions to counter earlier methodological problems. The digital environment allows for simplified and accelerated development of user-tests, alongside more accurate means of data collection (Montello, 2002). For example, whilst Phillips et al. (1975) relied on a series of written questions to assess the legibility of different relief depictions, Irvankoski (2012) and Popelka & Brychtova (2013) were able to utilise eye tracking and automated recordings of user response times – both of which helped reinforce the validity of their conclusions. A more detailed overview of such research approaches is given in the following section. In light of this progress, the new wave of map-design research which has emerged in the last two decades has thusly been characterised as "cognitive-digital" by Zyszkowska (2015). It is within this emergent sub-discipline that this thesis may be positioned.

3.2 *Methodological Approaches in Usability Studies*

A variety of different approaches have been taken in order to assess the usability of cartographic products. Before discussing each of these in turn, it logically follows to first define exactly what is meant by 'usability', and therefore exactly what must be assessed when usability is measured. As noted earlier, usability may be measured in terms of effectiveness, efficiency, and user-satisfaction (often, user preference) (Bevan et al. 2015). Each of these parameters are elaborated in Figure 3-3. Work by Frøkjær et al. (2000) found a lack of correlation between these criteria; usability studies should therefore assess all three criteria in parallel, to ensure all aspects of usability are accounted for.

Effectiveness

Effectiveness is a measure of how well a user is able to achieve what they set out to achieve when using a product, measured in terms of accuracy and completeness.

Efficiency

A relative measure of how accurate and complete a user is in achieving their goals, versus the resources (e.g. task completion time or mental effort) expended in reaching them.

Satisfaction

A subjective measure of a how pleased a user is with the cartographic product, and in general whether they personally 'like' it or not.

Figure 3-3 *Elaboration of usability criteria, adapted from Nielsen (1993), Nivala (2005) and Frøkjær et al. (2000).*

How, then, may each of these criteria be measured, and usability assessed? Broadly speaking, usability research has taken one of two contrasting research approaches: empirical user testing, and expert inspection methods. Empirical research approaches – also termed 'testing with users' (Whitefield et al. 1991), and 'user testing' (Holzinger 2005), have been described as the most fundamental form of usability testing (Nielsen 1993; Nivala 2005;

Holzinger 2005). This form of usability testing involves using 'real users' as experimental participants, to gain first-hand data on usability (Nielsen 1993). In contrast, 'expert testing', also known as 'usability inspection' (Holzinger 2005), typically evaluates designs by those who created them or those knowledgeable in their field. However, the expert approach is less adept at capturing real-life design and usability issues (due to separation from the end user), is unable to identify unknown user needs, and is at risk of experimental bias (Nivala 2005; Holzinger 2005).

As such, empirical-research has formed the primary basis of the majority of map-design usability studies. This research has generally taken one of several strands, most frequently: thinking aloud, user observation, questionnaires and interviews, focus groups, performance measurement and data logging (Nielsen 1993). A brief summary of these approaches is given in Table 3-1, with more detailed elaboration thereafter.

Research Approach	**Overview**
Thinking aloud	Users continuously vocalise their thought processes as they use a system.
User observation	Observer takes notes on user interaction with a system either: a) during interaction, b) from filmed interaction.
Questionnaires and interviews	Users questioned on their use of the system, by text or vocally.
Focus groups	Multiple users provide vocal feedback on a system, whilst being questioned as a group.
Performance measurements	A user's performance (typically response time, accuracy and error) is measured as they complete system tasks.
Data logging	A user's interactions with a system are automatically recorded – e.g. eye movement tracking, keyboard logging, mouse tracking.

Table 3-1 *Research approaches in the fields of cognitive map-design and map-usability. Adapted from Nielsen (1994), Holzinger (2005) and Farkas (2013).*

'Thinking aloud' is a form of usability evaluation where a user is asked to continuously vocalise their thought processes, during a usability test (Ivory 2001). Theoretically, this gives researchers comprehensive and direct insight into how a user interacts with and uses a system, alongside any problems they encounter, and misconceptions they maintain (Nielsen, 1993). As test participants talk, data may be transcribed directly, or subsequently from video or sound recordings. However, whilst the technique provides rich qualitative data, this richness also complicates analysis, making quantitative efficiency measurements and design comparisons more difficult (Holzinger, 2005). Though quantitative information can be derived (e.g. Kulhavy et al., 1992), doing so adds complexity, uncertainty, and time to analysis. Additionally, forcing a user to 'think aloud' is an unnatural request, which may lead to similarly unnatural user behaviour (Nielsen, 1993).

'Simple user observation' is an alternative method which avoids this last issue. Instead of asking a user to vocalise their thoughts and actions, researchers instead take notes on user behaviour as unobtrusively as possible, to attempt to capture data which is representative of real-life behaviour (Holzinger, 2005). Though the method is simple, it is mainly suited for discovering major user-issues, rather than the effect of subtle design changes. Pure observation is therefore less useful for research into specific design choices (Farkas 2013), as is the purpose of this thesis. However, it can be useful in discovering previously unknown issues, which may not have arisen in targeted laboratory tests (Nielsen, 1993).

Rather than assessing usability by simple observation, a particular improvement of the new wave of map-design research is the proliferation of automated data logging technologies (Zyszkowska 2015). User logging methods are those which automatically record the interactions of a user with a particular system (Farkas, 2013). In many cases, this logging may occur after the release of a system, to gain real user feedback in-situ, and often over a long time period (Nielsen, 1993). However, Holzinger (2005) note that user-logging can also offer valuable research insights in laboratory settings. For example, data logging is especially adept at recoding user response times (e.g. Savage et al., 2004), making it useful when measuring the efficiency of a system. Similarly, where a user test involves a digital questionnaire, data-logging can be used to automatically record user responses for each question (e.g. Brügger, 2015), greatly simplifying the experimental process.

In a digital-cartography setting, logging of mouse-movements and keyboard strokes offer notable opportunities. Mouse metrics (e.g. clicks, scroll wheel zooms, drags etc.) can be recorded, to determine which map components a user interacts with, and in what form that interaction occurs (Manson et al. 2012). Further tools can also produce mouse-click heat maps and analyse mouse movement, to give additional insight into the use processes and preferences (Boer et al. 2012). Yet perhaps the form of logging which has gained most traction in contemporary map-design research is 'eye-tracking', in which the eye movements of a user are recorded by specialised hardware and software (e.g. Brügger 2015; Popelka & Brychtova 2013; Ooms et al. 2014). The general theory of these 'eye movement studies' is that a user will focus on regions of a map where they are attempting to gain information. If the eye movements of a user are recorded, it therefore follows that it is possible to determine from which areas of a map information was obtained – and likewise which areas remained unused (Steinke 1987; Montello 2002). Temporal eye tracking data can give further insight into the order in which different map-areas were viewed, and thus provide useful information on the efficiency of different map designs, and the cognitive processes of the user (Incoul et al. 2015).

Though eye-tracking techniques have provided valuable insight in a number of empirical map-design studies, the data they create is elaborate and demands complex mathematical analysis. Perhaps more problematically, eye-tracking is of limited use when studying interactive and dynamic cartographic products (Ooms et al., 2014). Given that these new technologies allow the information contained within a set region of a digital display to change, an eye gazing on a particular region of a screen will not always observe the same area of a map. Naturally, this makes traditional eye-tracking techniques unsuitable for interactive and dynamic displays (Liao et al. 2016). Though research by Ooms et al. (2014) developed a methodology to overcome this issue, by combining eye-tracking with other forms of user logging (mouse movements and keyboard strokes), a comprehensive and flawless solution remains lacking.

Data logging, especially of response times, is commonly used to support 'performance measurement', a research approach that aims to gather quantitative data on the performance of a user when completing tasks (Ivory 2001). The measured performance parameters depend upon the aims of the study, but – beyond response time – typically include measurements of response accuracy and error. Performance measurement is thusly fundamental when assessing usability, as it allows determination of effectiveness and accuracy (c.f. Figure 3-3). However, Nielsen (1993) stresses the importance of ensuring any measured performance parameters directly support each study's stated research aims. Likewise, ensuring that any performance tasks are as realistic and representative as possible is crucial (Farkas, 2013).

Whereas the preceding research methods provide direct insight into the usability of a system, questionnaires and interviews (e.g. Schobesberger & Patterson 2008; Bleisch & Dykes 2008) provide indirect insight, through the filter of a user's opinions (Holzinger, 2005). As such, they are especially useful for gauging user preference and satisfaction (Nielsen, 1993). Holzinger (2005) also notes that statistical information can be easily derived from both questionnaires and interviews, though a high number of participants – generally at least 30 – is needed to produce statistically valid and robust conclusions. Care must be taken to avoid designing questions which influence the user and bias results (Farkas, 2013). Nielsen (1993) also stresses that questionnaires must undergo pilot testing, partly to avoid said bias, but also to ensure all questions are easily understandable. However, even when questionnaires are well designed, their data must be scrutinised given its inherent subjectivity: what users think they want, may not actually equate to what they *need* (Nielsen, 1993; Holzinger, 2005).

A natural extension of single-user questionnaires and interviews is the focus group (e.g. Sherwin & Bartle, 2012), whereby a number of typical users are brought together in feedback sessions (Nielsen, 1993). Farkas (2013) suggests that the aim therein is to improve the richness and quality of user feedback, by stimulating debate and provoking discussion. However, focus groups suffer from the same subjectivity issues as questionnaires and interviews, whilst considerable skill as a moderator may be required in order to encourage appropriate and useful feedback (Nivala, 2005). Further, running a number of focus group sessions is necessary to produce representative results, as the outcome of a single session may be biased (Nielsen, 1993).

Whilst the preceding research approaches have been grouped into seven distinct categories, it must be stressed that a combination of methods is often advisable, and often apparent in the literature, when usability is assessed. For example – while efficiency may be best judged using performance measurement, user satisfaction is likely better evaluated using questionnaires or interviews. Gaining a comprehensive overview of the usability of an entity therefore demands a multi-faceted research approach. The exact research methods utilised in this thesis, alongside further justification of their choice, are given in Chapter 5.

3.3 *Empirical Cycling-Map Research*

Despite the long history of cycling maps (Akerman, 2002), empirical research centred on their design, not least their depictions of elevation, is notable by its scarcity. However, that is not to say that academic cartography has totally bypassed the topic. Indeed, recent work by Wessel & Widener (2015) aimed to design an 'objective' urban cycling map for the city of Cincinnati, which communicated information pertinent to all forms of cyclist

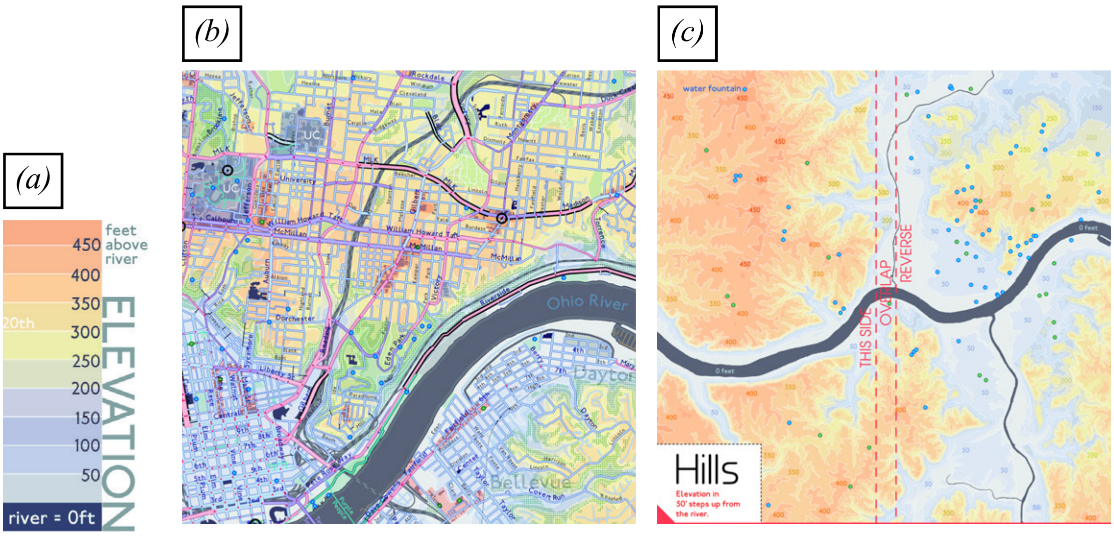


Figure 3-4 *Elevation depictions in the Cincinnati bike-map designed by Wessel and Widener (2015), including: (a) elevation legend, (b) elevation with street network, and (c) elevation without street network (note the different scales).*

– from competitive riders to everyday commuters. The authors attempted to achieve this by encoding only that information, including terrain, which they considered universally important. The resulting map communicated elevation information through hypsometric tinting, without contour lines (Figure 3-4). However, their research lacks a thorough user evaluation of the design that was developed. No formal user-testing or user-surveys of their map were conducted, with the authors instead relying on self-submitted feedback forms to gauge user opinion. Their broadly positive review of the map's design, and style of elevation depiction, is thusly purely anecdotal.

From a route-planning perspective, Dickinson (2012) developed an experimental cycle route-planner that placed an emphasis on communicating cycling experience, rather than simply providing the shortest or fastest route. Thus, tool enabled users to plan a route based on a combination of street conditions, physical surroundings, traffic speed, road surface, and elevation information. Two forms of elevation depiction were assessed in this

Figure 3-5 *Alternative elevation depictions developed by Dickinson (2012) for his experiential cycle route-planner, using: (a) small multiples of elevation profiles for different routes, and (b) trapezoid/arrow depictions.*

study. The first utilised a series of elevation route profiles, displayed simultaneously (i.e. small multiples) in the perimap area. Each profile is effectively a graph of elevation along a particular route (Figure 3-5a). The alternative visualisation developed by Dickinson (2012) used arrow symbolisation (Figure 3-5b): the placement of the arrows represented hill direction and location, whilst the steepness of each hill was denoted by the width of each arrow's base. Dickinson assessed the user preferences for these designs through a combination of thinking aloud and interviews, using thirteen participants. The majority of those interviewed preferred the route profiles to arrow depictions, finding them clearer, more intuitive, and more easily comparable (Dickinson, 2012). However, while the study briefly considered user preference, efficiency and effectiveness were not assessed, making the study a poor measure of the usability of the route planner, and elevation depictions, as a whole. Moreover, the design of elevation depictions formed only a tiny aspect of the study, whose primary purpose was to assess the usefulness of experiential cycle route-planners in general.

A more rigorous user-study was conducted by Sherwin and Bartle (2012), who conducted focus-group based research into the usability of four different 'styles' of cycle mapping, using 29 participants. However, rather than considering the style of cartographic visualisation, the primary focus was upon which information cyclists found most useful. Elevation depictions were briefly considered, but only in a simplistic, binary sense – that is, whether or not they should be included or omitted. All except one focus-group participant thusly concluded that contour lines were 'essential', with the dissenter preferring instead arrow depictions (Sherwin and Bartle, 2012). However, this result is questionable, given that participants were shown no forms of elevation depiction other than contour lines. Instead, it would be more reasonable to conclude that *elevation depictions* were essential. Moreover, as with the work of Dickinson (2012), whilst the study ostensibly assesses usability, it fails to take into account the efficiency or the effectiveness of each design. Following the aforementioned results of Frøkjær et al. (2000), it therefore fails to provide a comprehensive representation of the relative usability of each cycle mapping 'style'.

The preceding literature focussed on the design of cycling maps in general, with varying degrees of user-testing, and varying emphasis on elevation visualisations. As yet, only the MSc thesis of Brügger (2015), later published in a revised form in the journal Cartography and Geographic Information Science (Brügger et al., 2016), has specifically assessed elevation depictions in cycling maps with an empirical user study. The usability of three elevation visualisations were studied: elevation route profiles, arrow symbolisation, and

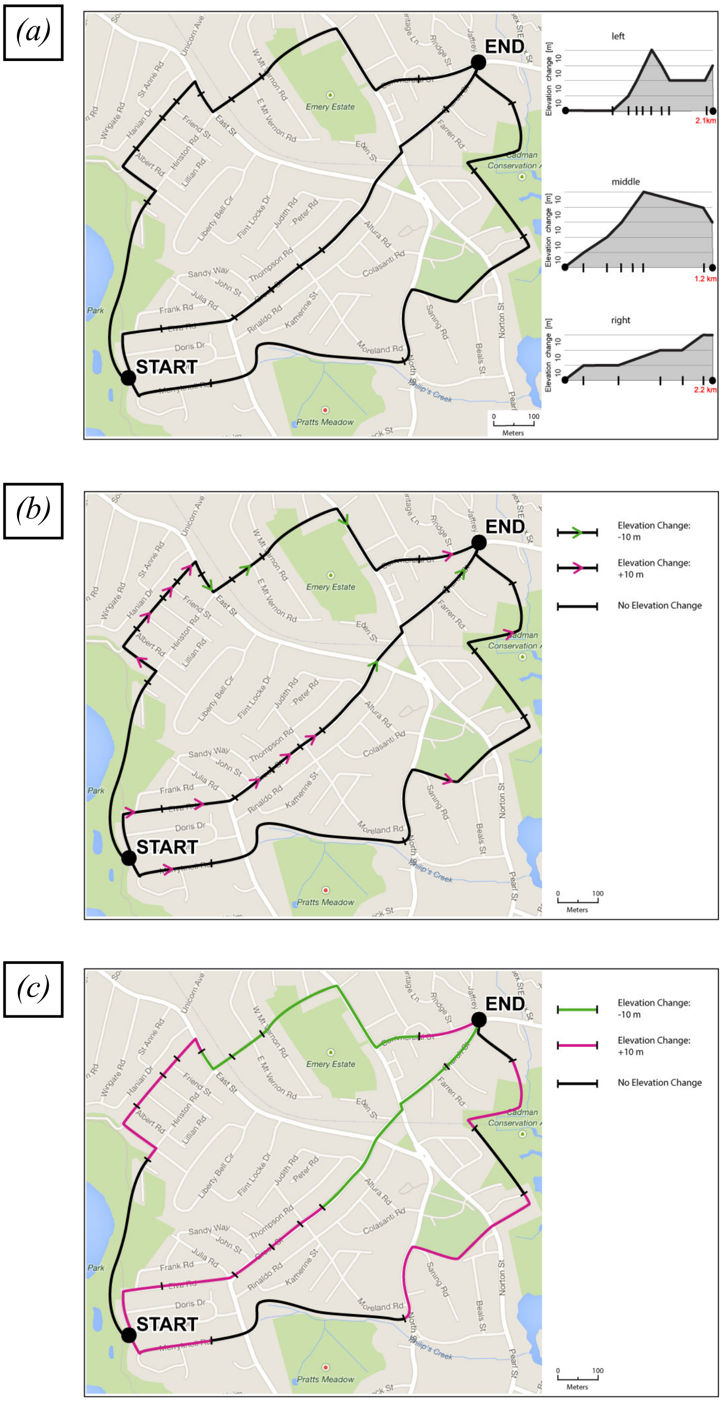


Figure 3-6 *The three forms of elevation depictions used in the comparative empirical study of Brügger (2015) – (a) elevation profiles, (b) arrow depictions, and (c) colour representation.*

colour-coding (Figure 3-6). In all instances, the representations were linear; they represented elevation only along distinct paths (Brügger, 2015). Each style was assessed in terms of efficiency, effectiveness and user-preference, for two typical use tasks – finding the steepest slopes ("slope identification") and the biggest elevation changes ("height detection") (Brügger, 2015).

A mixed-method research approach was taken during the study. The usability of each depiction was assessed using an online survey, which users filled in as they completed slope identification and height detection tasks. Automatic logging was used to record response time (efficiency), and the accuracy of response was calculated dependent on the given answers (effectiveness). Users were also asked about their general map preferences throughout the task (preference). In addition, eye-tracking was also used to ascertain how users perceived and dealt with different elevation depictions. Brügger et al. (2016) used a within-subject design, meaning that every participant was exposed to every combination of design and task. When assessing slope identification, the study found no significant difference between each representation type. However, users were more accurately and more quickly able to perform height detection tasks with arrow and colour elevation depictions ('intrinsic' depictions – within the map), as opposed to elevation route profiles ('extrinsic' depictions – separate from the map). On the basis of the eye-movement data, the authors suggest that this may be due to the forced split in visual attention demanded by separating the elevation information from the map itself. Thus, it would be beneficial to study intrinsic route profiles (c.f. those in Figure 3-16), to determine if this is indeed the case. Likewise, all the elevation depictions studied by Brügger (2015) are static and two-dimensional – further research into dynamic and three-dimensional depictions would therefore help to reinforce or challenge the results of the study. Brügger (2015) also found that, counter-intuitively, users tended to prefer those depictions which were least efficient and least effective (i.e. the elevation profile).

However, the results of Brügger (2015) may be questioned partially on the basis of the study's methodology. In particular, a lack of variation in the geographic areas selected for the study has the potential to have influenced its results. At the task level (e.g. selecting the highest marker) the geographic area was kept constant, while the visualisation type was changed. This design is summarised in Table 3-2, which gives an overview of the study for a theoretical participant.

CITIES	**TASKS**		
	Route Choice *(which route would you take?)*	**Highest Point** *(which of these three points is highest?)*	**Steepest Slope** (*which slope is steepest?)*
Boston	V1, V2, V3	V1, V2, V3	V1, V2, V3
Perth	V1, V2, V3	V1, V2, V3	V1, V2, V3

V1 = Arrow visualisation, V2 = Elevation profile visualisation, V3 = Colour visualisation

Table 3-2 *An overview of the questionnaire design utilised in Brügger (2015). The table shows the form of questions undertaken by a theorized participant in the study.*

The intention was to ensure that only the effects of changes in visualisation type, and not environment, were the cause for any recorded variations in efficiency or effectiveness. However, it is plausible that a participant could recognise the study area on second or third viewing. Thus, when asked for the second time which marker was highest, the user may already know the answer. Though Brügger (2015) attempted to counter this by rotating the maps and including two cities in the study, some degree of recognition is still highly likely. This casts doubt on the reliability of the study's results. Future research should aim to minimise this learning effect to a greater degree, if possible.

3.4 *Existing Elevation Depictions in Cycling Maps and Cycle route-planners*

Given the paucity of research which has focussed explicitly on elevation and terrain depictions in cycling maps, it is logical to present an overview of the different styles of elevation depiction present in contemporary cycling maps and cycle route-planners. In line with the aim of this thesis, these depictions are separated into those which are two- and three-dimensional. In her MSc thesis, Brügger (2015) comprehensively reviewed cycling map depictions. What follows is therefore only a brief overview, with an additional focus on interactive 3D depictions, to avoid repetition.

3.4.1 *2D Elevation Depictions*

Given that, at present, the majority of cycling maps and cycle route-planners are wholly two-dimensional, so too are the majority of their elevation depictions. Beyond pure dimensionality, Brügger (2015) further made a distinction between those depictions which are continuous and discontinuous; this distinction is used as the basis for the following review.

3.4.2 *Continuous 2D Elevation Depictions*

Continuous elevation depictions are those which cover the entirety of the mapped geographic area. Arguably the most widely adopted method which falls under this category is the use of contour lines – lines of constant altitude that provide an impression of overall terrain (Figure 3-7). Though it is possible to use contours alone, they are often combined with additional visualisation techniques. Hill-shading (Figure 3-8) is commonly used as a layer beneath contour lines (though it may also be used independently), to provide an impression of realism (Fabrikant & Wilkening, 2011). In essence, the aim of hill-shading is to generate a 3D impression of terrain within the mind of the map-reader (Collier et al. 2003). An alternative technique is to utilise hypsometric tinting, either independently, or in combination with contour lines and/or hill shading. Hypsometric tints, also termed 'layer' and 'elevation' tints, visualise elevation through colour, most frequently either as a series of distinct bands, or as a graduated colour ramp (Patterson & Jenny 2011).

Figure 3-7 *A sample cycle map for the city of Bristol, using contour lines to depict terrain. Suggested cycling routes are shown in green. (Sherwin and Bartle, 2012).*

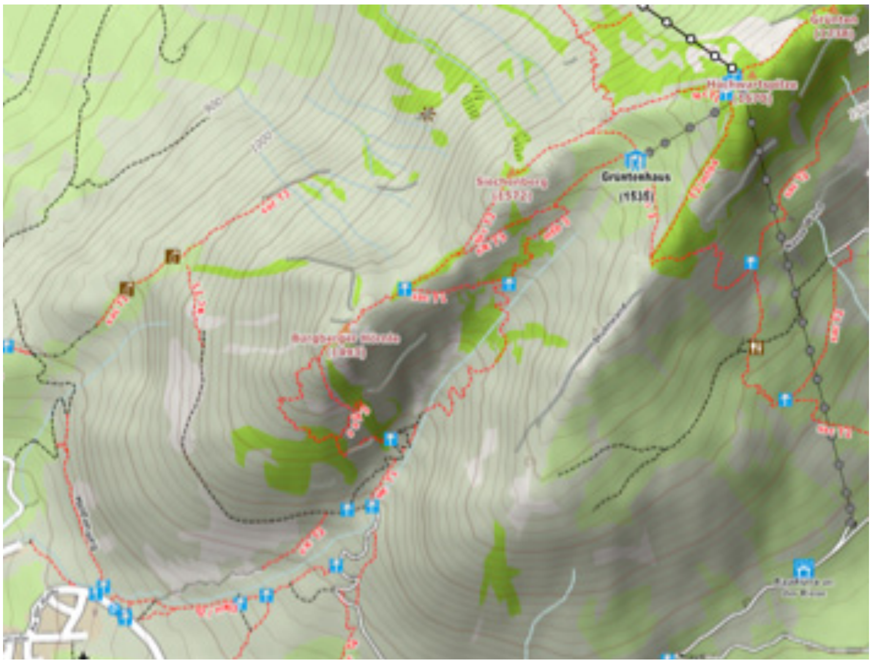


Figure 3-8 *Hill-shading used in combination with contour lines to depict terrain on a '4UMap' map of mountain-bike trails in Bavaria, Germany (4UMap, 2012)*

Yet whilst all such depictions have been used on cycle maps, they are not without issues. In urban contexts, the density of human development is likely to obscure continuous elevation information. This was a problem, for example, in the Cincinnati bike map of Wessel and Widener (2015), which featured banded hypsometric tinting. The authors attempted to counter the issue by including an additional terrain map – free of roads and other human development (Figure 3-4) – but this is an evidently cumbersome and unworkable solution. Patterson & Jenny (2011) further note that some users are likely to mistake hypsometric tints for depictions of other landscape features, such as forests. Shaded relief also has noted issues, with some research suggesting the added realism reduces the accuracy and efficiency of spatial decision making (Wilkening and Fabrikant, 2011).

3.4.3 *Discontinuous 2D Elevation Depictions*

In light of the issues with continuous depictions, Brügger (2015) suggested that discontinuous elevation representations – in particular 'linear' elevation depictions – provide a superior option for cycling maps and cycle route-planners. In these visualisations, terrain information is presented only for the route of a path, or multiple potential paths. Because a smaller proportion of the map area is used to display elevation, the problems of cluttering and obstruction associated with continuous depictions are avoided (Brügger, 2015).

Linear depictions may also more suited to displaying slope rather than absolute height (Brügger, 2015) – this in itself may be beneficial for cyclists. Though research into cycling motivators has broadly demonstrated the importance of terrain, many studies point to 'steepness' as a more important factor than elevation (Hochmair 2004; Su et al. 2010; Winters et al. 2011; Huffman 2009). This suggests that cycle maps and route planners should focus on mapping slope rather than absolute height. Though Wilkening and Fabrikant (2011) note that all elevation depictions can be used to infer the steepness of a slope, some do so with much less cognitive effort than others. For instance, the cognitive load for a user to derive slope from a contoured map is likely much higher than for an elevation profile (a form of linear elevation depiction), where slope is clearly visible on first glance. For this reason, from a cycling perspective, the attraction of linear elevation depictions is clear. Brügger (2015) further suggests that linear depictions are in essence a form of task specific generalisation and simplification – displaying only the information which is of interest to a cyclist – and as such a design choice which follows good cartographic practice.

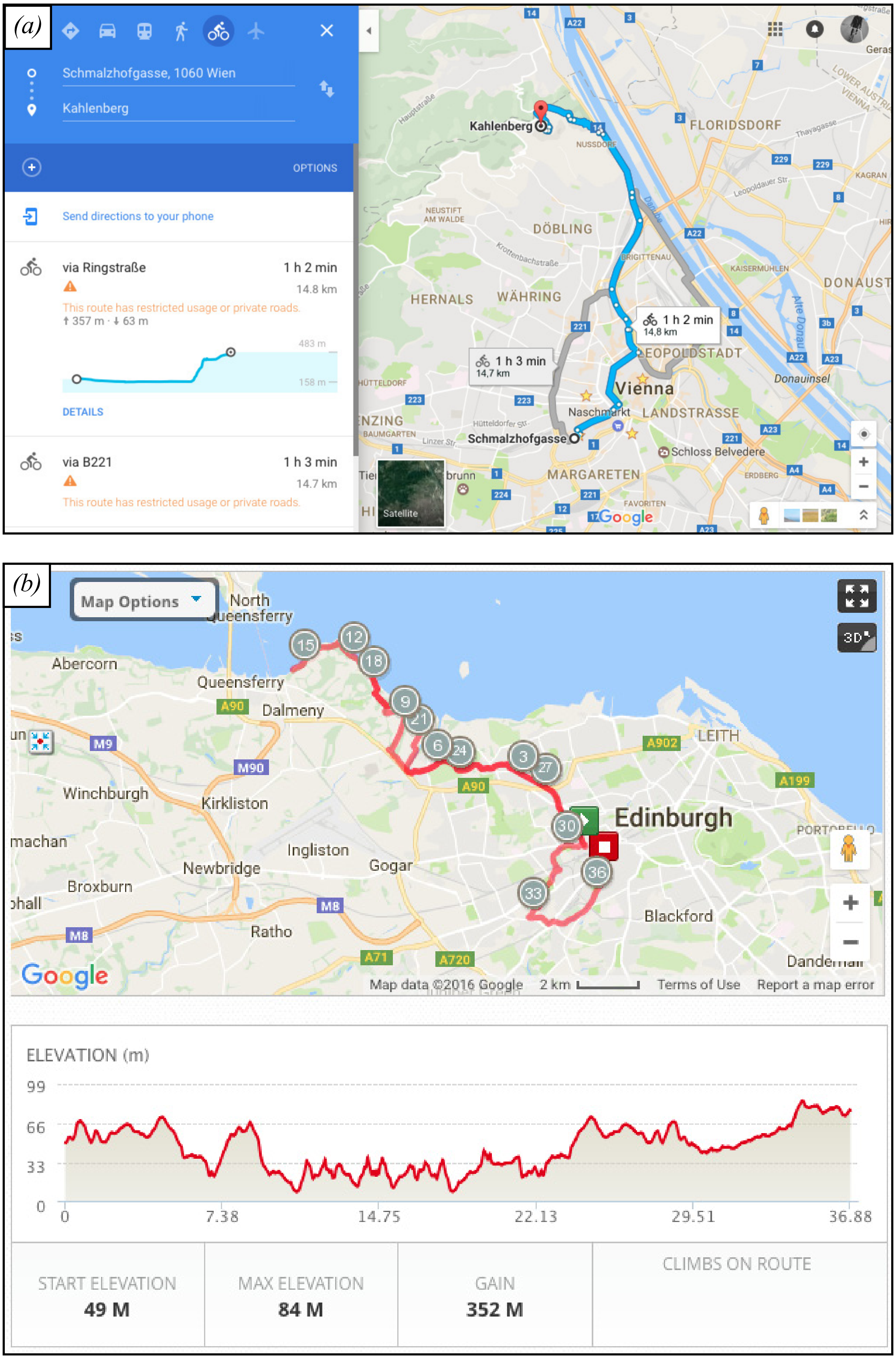


Figure 3-9 *Elevation profiles as featured in two popular online cycle route-planners: (a) Google Maps, and (b) 'MapMyRide'. (Google, 2016; MapMyFitness, 2016)*

As noted, elevation profiles – also termed hill profiles and route profiles – are one such linear depiction; they represent an effective cross-section of the terrain over which a route traverses. Elevation profiles are utilised in a number of online cycle route-planners (Figure 3-9). The profiles provide a detailed yet intuitive representation of route terrain, and are especially adept at communicating the slopes on a path (Brügger, 2015). However, as they are separated from the map itself, users may have difficulty relating points on the profile to points on the route (Dickinson, 2012). In order to counter this, some cycle route-planners interactively link elevation profiles to mapped paths. For example, both Google Maps and MapMyRide allow a user to run their cursor along the elevation profile while simultaneously viewing the corresponding point on the mapped route.

To avoid the separation of map and elevation information, and the visual attention split that results, Huffman (2009) suggested encoding elevation information along the route path, on the map itself. In line with the aforementioned research, the author suggests that mapping slope is most useful to cyclists, although simultaneous encoding of both elevation and slope is possible with line thickness (Huffman, 2009). Most commonly, slope has been encoded through line thickness, colour, and arrow symbology (Brügger, 2015). In his work, Huffman (2009) suggested that line thickness offers a number of advantages in this

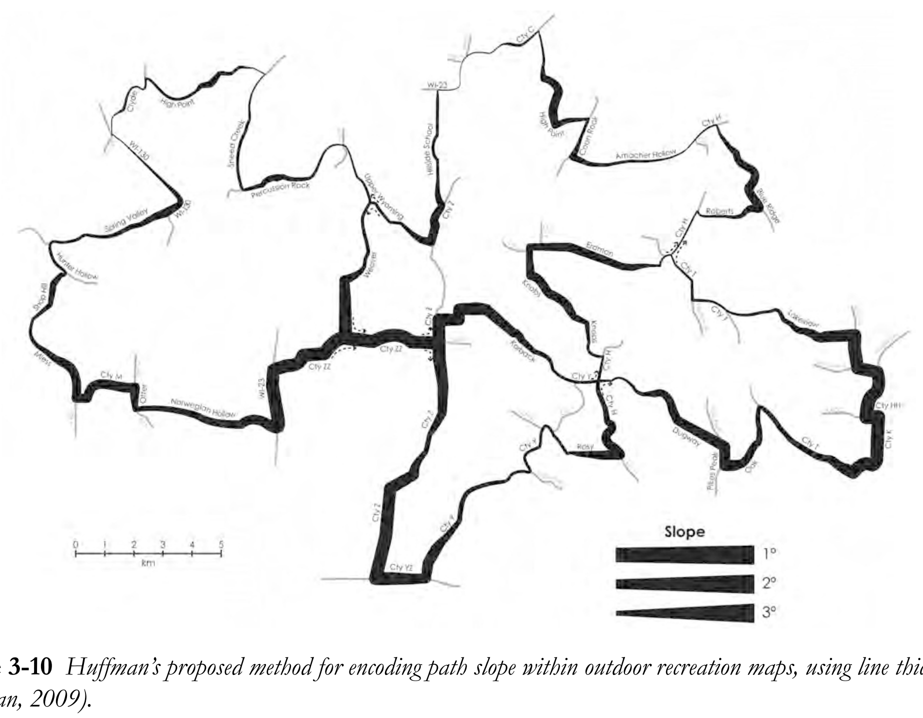


Figure 3-10 *Huffman's proposed method for encoding path slope within outdoor recreation maps, using line thickness (Huffman, 2009).*

respect. Changing line thickness (Figure 3-10) should allow the map reader to much more easily quantify changes in elevation or slope than arguably more subtle changes in colour, saturation or hue (Huffman, 2009), although no empirical evidence is given to support these statements. Similarly, there should be no difference in interpretation of line thickness under different lighting conditions, whereas this may be an issue where colour is used as the encoding variable (Huffman, 2009). However, Brügger et al (2016) make the point that variable line thicknesses may be misinterpreted as changing road types or conditions.

In spite of the noted issues with colour encoding, a number of examples exist where colour hues are used to encode slope in a cycling context. In an academic context, Su et al. (2010) used colour as a means of communicating slope in their cycling maps of Vancouver, Canada. Other examples outwith academia demonstrate the existence of this technique in the mainstream (Figure 3-11). Unlike variable line thicknesses, colour variations have the advantage that they always require the same amount of space on a map – whereas thicker lines may lead to occlusion and obstruction (Huffman, 2009). However, Brügger (2015) makes the point that where colour is used as the encoding variable, street-names may necessarily need to be displaced; this may also lead to problems of occlusion, especially in dense urban areas. Further, although Su et al. (2010) suggested that colour encoding of slope would allow the map reader to quickly and easily assess the overall hilliness of a route, the empirical results of Brügger (2015) suggest that arrows are much more efficient and effective in this respect.

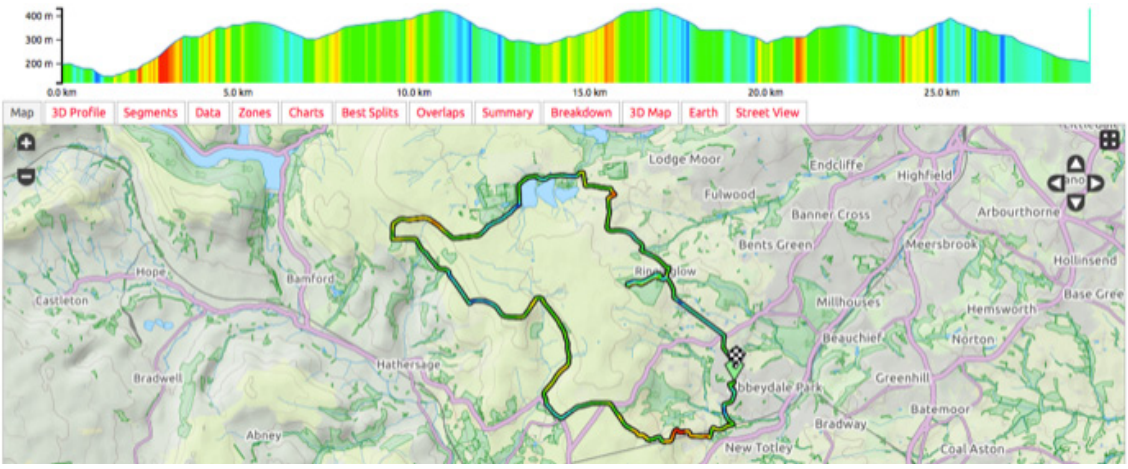


Figure 3-11 *In tandem with an elevation profile, the 'VeloViewer' website visualises path slope on the mapped path itself using colour – red indicates a steep upward slope, blue a steep downward slope, and green a flat surface – with varying shades inbetween (VeloViewer, 2016).*

Indeed, a number of cycle maps and cycle route-planners use arrows to encode slope. Brügger et al. (2016) note that arrow shape and orientation can be modified to encode a wide variety of variables, including direction, distance, and slope. Recent cycling examples include the official Vancouver and Seattle cycling maps (Figure 3-12). Huffman (2009) further showed the flexibility of arrows (Figure 3-13) , demonstrating how they may be used to simultaneously communicate uphill and downhill (arrow direction) and slope (arrow size). This was an approach taken by Dickinson (2012) when designing an experiential cycle route-planner for Vancouver (Figure 3-5b). However, one of the benefits of arrows – their flexibility for encoding a wide variety of information – is also their downfall. Brügger et al. (2016) suggest that their wide-ranging applicability makes them less intuitive. Upon seeing an arrow, a map-reader is unlikely to know whether it represents slope, direction, or some other variable, without first consulting the legend. Indeed, this is an issue with the Seattle and

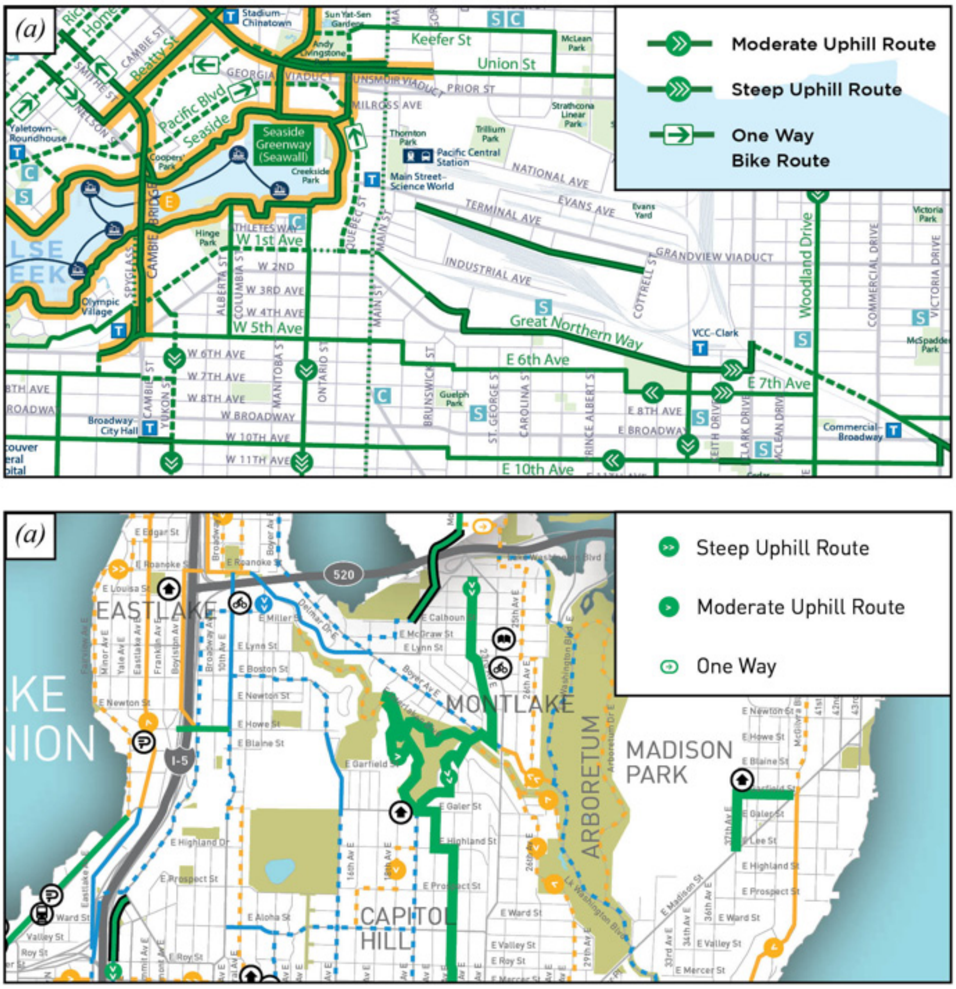


Figure 3-12 *Arrow based depictions of road steepness, alongside associated legend explanations, in the official cycling maps of the cities of (a) Vancouver, and (b) Seattle. (City of Vancouver 2016; Seattle Department of Transportation 2016).*

Figure 3-13 *Some possibilities for arrow symbolisation. Arrows indicate downward slope direction, with steepness encoded by: (a) arrow size, (b) colour saturation, and (c) arrow size and colour saturation. Based on the work of Huffman (2009).*

Vancouver cycling maps in Figure 3-12, both of which feature three types of arrow – one for moderate uphills, one for steep uphills, and one for one-way streets.

3.4.4 *3D Elevation Depictions*

The advent of 3D cartography has brought with it a myriad different forms of 3D mapping. For clarification, the term '3D' in this thesis refers to all forms of 'pseudo-3D' depictions. Following the definition of Buchroithner and Knust (2013), these are monoscopic visualisations which may be displayed on traditional planar media. Such visualisations rely on a series of monoscopic depth cues – such as illumination, shadow, texture gradients and aerial perspective – to create a mental impression of a faux 3D scene (St John et al. 2001). They may be contrasted with 'real-3D' or 'true-3D' depictions, for which two distinct images are stereoscopically presented to the viewer, using specialised hardware.

3.4.5 *Photorealistic 3D Visualisations*

Pseudo 3D cartographic depictions are themselves incredibly diverse. In what Popelka & Dolezalova (2015) term 'photorealistic 3D visualisations', the goal is imitation of the real world. Such depictions frequently use aerial imagery draped over 3D digital terrain models,

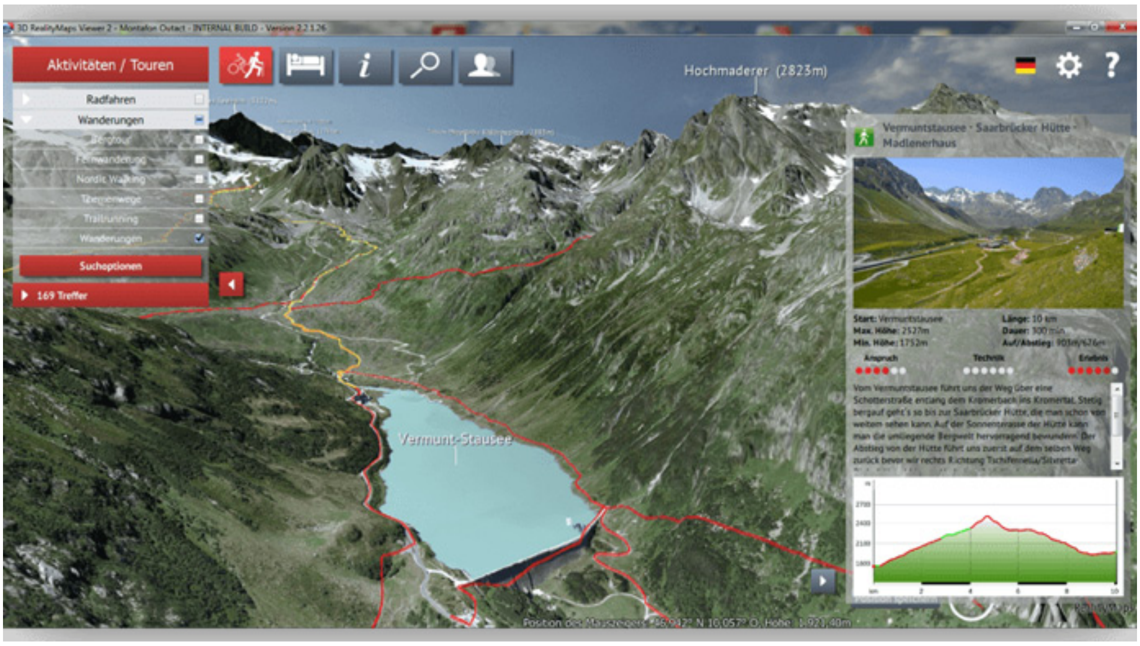


Figure 3-14 *An example of a recreational route displayed in an interactive and photo-realistic 3D environment. The environment is comprised of a 3D digital terrain model, with aerial imagery draped over the surface, and was developed by 'Reality Maps' (Reality Maps, 2016).*

Figure 3-15 *A screenshot from the route planning function of the 'Maps 3D PRO' mobile application. (Movingworld GMBH 2016).*

to provide more realistic spatial depictions (e.g. Figure 3-14). However, the majority of cycle route-planners which provide photorealistic 3D visualisations do so not for the purposes of aiding route planning, but instead as a pre-or post-ride aesthetic visualisation. This purpose is seen, for example, with tools such as 'Relive.cc', which automatically converts GPS cycling tracks into 3D flyover videos (Relive 2016).

In some instances, a 3D digital terrain model is used without an image overlay, though the aim generally remains added realism (e.g. Schobesberger & Patterson, 2008). An example in this instance is the 'Maps 3D Pro' mobile application, which places a greater emphasis on route planning than the aforementioned examples. Unlike the tools in Figure 3-14, this application allows users to plan cycling routes directly within a 3D map (Figure 3-15). Some empirical studies have attempted to assess the usability and performance of such depictions, with conflicting conclusions (Section 3.5). A key criticism in the context of this thesis, is that while such visualisations may seem aesthetically pleasing, they are arguably more suited to rural environments with low levels of urban development. Highly developed urban areas may instead be better suited to non-photorealistic 3D visualisations.

3.4.6 *Non-Photorealistic 3D Visualisations*

In contrast, 'non-photorealistic 3D visualisations' may include additional 3D objects placed atop a 2D (or 3D) basemap. The goal in this case is not necessarily added realism, rather the use of an additional dimension to improve cartographic communication (Rautenbach et al. 2015). Notably, there has been recent growth in the availability of intrinsic 3D elevation

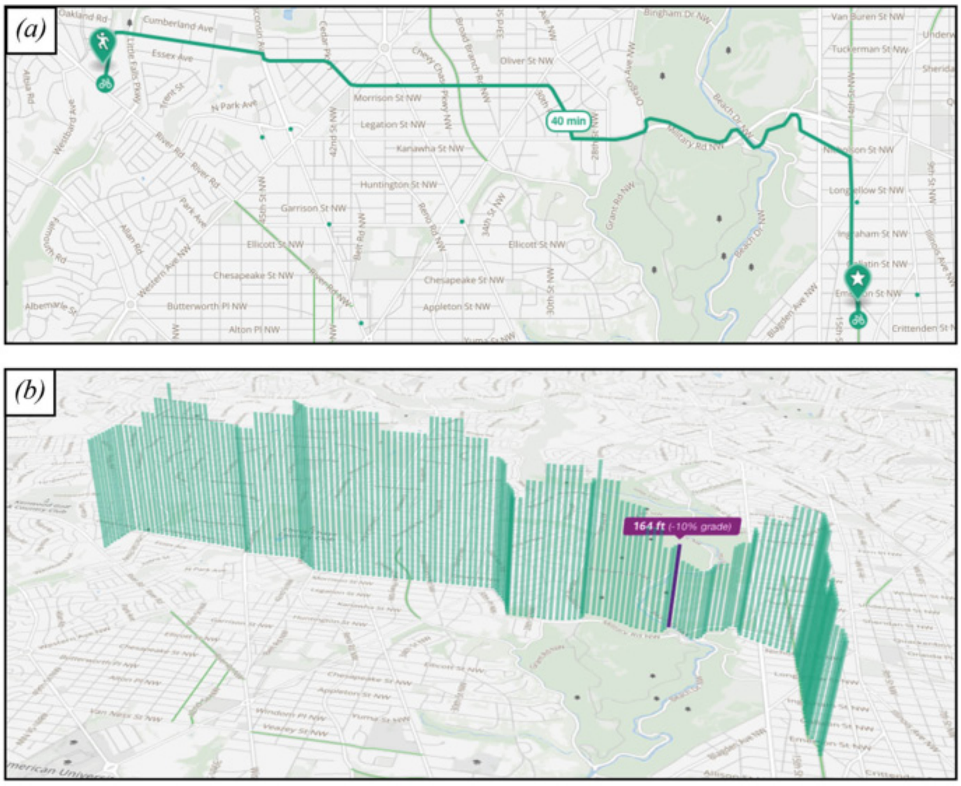


Figure 3-16 *The conceptual route planning interface developed by MapBox for the Washington Bike Share scheme, featuring (a) the 2D route planner, and (b) the interactive 3D elevation profile for the route shown in (a). (Liu, 2015).*

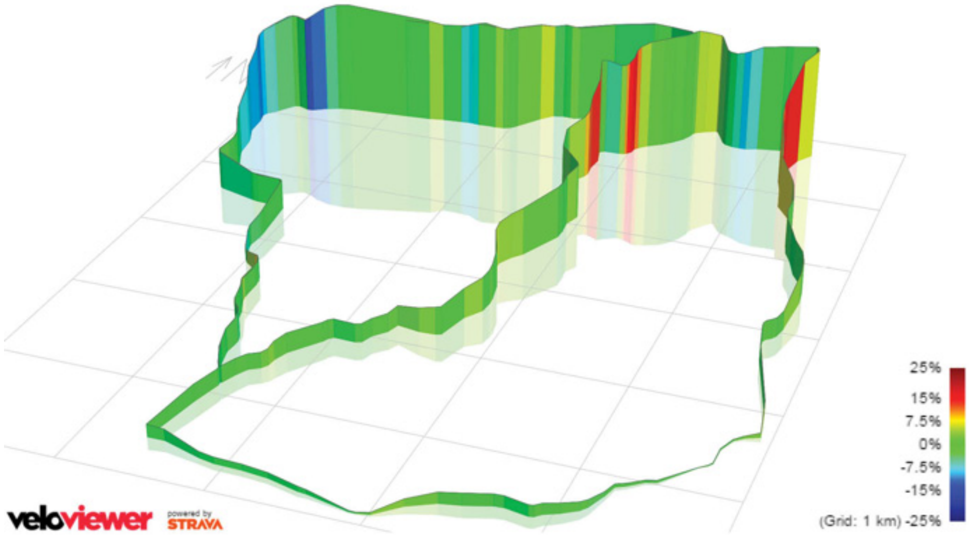


Figure 3-17 *The interactive 3D route profile developed by VeloViewer for visualising cycling routes. The varying colours relate to changing slopes of the route path. (VeloViewer, 2016)*

profiles into cycle route-planning tools. These depictions bring route profiles onto the map's surface, extruding them to the height of the terrain beneath. One of the earliest interactive examples was produced by MapBox for the Washington bike-share scheme (Liu 2015), allowing users to plan a route whilst switching between a 2D overlay and a 3D route profile (Figure 3-16). An additional prominent example is 'VeloViewer', which further highlights steep slope sections through colour coding (Figure 3-17). Though these depictions offer a novel solution – especially for densely populated urban areas – their usability has not yet been empirically tested.

3.5 *2D vs. 3D Visualisations – Existing Comparative Studies*

The cartographic examples included in Section 3.4 highlight the potential of 3D visualisations for communicating elevation information. Though research comparing the relative usability of these visualisation forms with their 2D counterparts has not yet been carried out for cycling maps or route planners, similar work has been conducted for a small number of other cartographic use scenarios. The results of these studies are not necessarily applicable to cycling scenarios, given the findings of Liao et al. (2016), Schobesberger & Patterson (2008) and Savage et al. (2004), amongst others, which suggest the relative benefits of 2D and 3D spatial visualisations are heavily dependent on use context and task. Nonetheless, they merit further review – both to place this thesis in a wider context, and to help inform its research objectives and methodology.

Interest in potential 2D–3D usability differences has grown concurrent with a rise in the proliferation and ubiquity of 3D cartographic depictions, coupled with a growth in digital cartography as a whole (Shepherd 2007; Herman & Stachon 2016; Incoul et al. 2015; Kalt 2015). The increasing processing-power and capabilities of mobile and web technology further helped the spread of these new forms of cartographic visualisation (Incoul et al. 2015). However, the results of subsequent research into their usability have been contradictory, whilst the methods used to derive them are also worthy of scrutiny.

The potential benefits of 3D cartography may at first glance seem numerous. Depicting the third dimension – i.e. relief or elevation – on a 2D map is a problem which has challenged cartographers for generations, and which continues to cause headaches (Collier et al. 2003). At its most fundamental level, 3D cartography permits an additional dimension of visualisation, and so *should* help solve this problem. Especially in photo-realistic 3D visualisations, the assumption has been that the extra dimension breeds intuitivism, allowing users to more easily make spatial judgments and obtain elevation information from the visualisation (Niedomysl et al., 2013). The results of Schobsberger and Patterson (2008) seem to support this. The authors studied the communication efficiency of printed 2D and 3D perspective hiking-trail maps at a national park in Utah, using a combination of questionnaires and interviews with 185 hikers. The results of the study suggested that 3D depictions were marginally better at communicating spatial information – topography, distances and environment – and were significantly better at allowing the hiker to locate their position. 3D perspective maps were also found to be generally more appealing for hikers. (Schobsberger and Patterson, 2008).

However, the general user preference for 3D depictions described in Schobsberger and Patterson (2008) is not necessarily a good indicator of their true usability, given the issue of 'naïve realism', proposed by Smallman and John (2005). This concept theorises that users tend to prefer 'realistic' and seemingly intuitive visualisations, that are in reality less effective and efficient. Indeed, further research has highlighted a series of key problems associated with 3D cartographic depictions, namely the occlusion of objects, distorted and non-constant scales, and the perceived foreshortening and lengthening of objects or surfaces, depending on the viewing angle (Niedomysl et al. 2013). Though interactivity may tackle some of these issues – for example, by allowing a user to re-orientate the map to counter occlusion – such problems cannot be ignored, especially in static media (Niedmosyl et al., 2013).

Moreover, further comparative empirical research has drawn contrasting conclusions to those of Schobsberger and Patterson (2008). Savage et al. (2004) evaluated the performance of 2D topographic maps relative to their 3D perspective equivalents. The study assessed 68 participants on a series of questions related to determining location, elevation, distance and terrain shape, in terms of accuracy and efficiency (response time). 2D maps were found to be superior on both measures for all tasks, aside from those involving terrain, for which 2D and 3D depictions were found to have an equivalent performance. Similar conclusions were reached by (Bleisch & Dykes 2006), who evaluated the usefulness of photorealistic 3D visualisations for planning hiking routes in the alps. Using comparable research methods, the authors found that 3D maps were a 'poor solution' for extracting exact terrain information. In the study, test participants were rarely able to correctly determine the steepest slopes of a hiking route, or the length of the route as a whole. However, the study did show that the 3D visualisation gave users a comprehensive spatial overview of the test region. This supports the findings of Schobsberger and Patterson (2008) that 3D maps are better able to help a user identify their location – a task that would be aided by comprehensive spatial overview.

Whilst Savage et al. (2004) and Bleisch & Dykes (2006) used a combination of questionnaires and interviews in their research, Popelka & Brychtova (2013) additionally

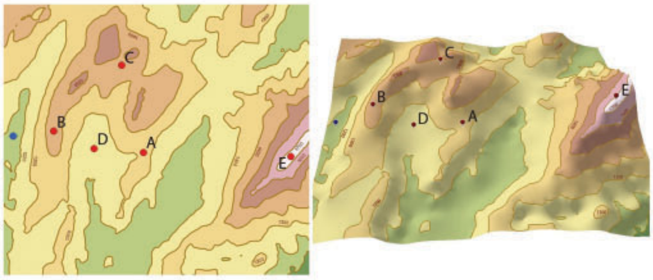


Figure 3-18 *The two forms of map-stimulus used in the study of Popelka and Brychtova (2013) – hypsometric shading (left) and 3D perspective-view with hill-shading (right).*

utilised eye-tracking methods to study user perception of 2D and 3D maps. The authors compared hypsometrically shaded 2D contour maps, and their 3D perspective equivalents (Figure 3-18). Initial eye-tracking data from this study found no perceptual differences between the two visualisation forms. However, further analysis suggested this result was influenced by the simultaneous depiction of both 3D and 2D map stimuli. An alternative methodology, using re-arranged stimuli, subsequently found significant differences in only one eye-movement metric ('scanpath length' – the total length of pixels scanned when looking at a map). The results suggest that the form of 3D perspective map studied is slightly more efficient than its 2D counterpart at communicating terrain information. However, the lack of confirmation of this result in any other eye-tracking metric (e.g. point fixation duration) casts doubt on the conclusion.

A key piece of research was conducted by Wilkening and Fabrikant (2011), who in part assessed how 'realism' affected three-dimensional map-based decision making. Though the authors did not use pseudo-3D or perspective maps in their research, they did study two forms of hill-shading, which were designed to give an impression of '3D-realism'. User performance in a series of decision-making tasks for these maps was assessed relative to simpler, more abstracted map forms (contour lines only, and contour lines with coloured slope classes). The results of the study echoed the aforementioned research into 'naive cartography'. Whilst users preferred realistic hill-shaded maps, that realism also had a detrimental impact on accurate decision making (Wilkening and Fabrikant, 2011). However, this result can not necessarily be expanded to conclude that all 3D maps are detrimental to decision making; such a conclusion would deny the existence of the range of 3D depictions noted in Section 3.4. In particular, the 3D elevation models from the likes of Liu (2015) are more abstracted and simplified in their nature, and inherently unrealistic. The usability of these kinds of geometric 3D abstractions should therefore be assessed, to determine if they suffer the same issues as 3D visualisations whose goal is enhanced realism.

Crucially, all the preceding studies used static rather than interactive or dynamic 3D terrain visualisations. This raises the question, do the results of research into static 3D representations also apply to their interactive 3D counterparts? Only a handful of empirical research has thus far studied the latter. Wilkening & Fabrikant (2013) studied user interaction with an interactive 3D geo-browser (Google Earth), by asking test participants to perform a variety of 3D cartometric tasks, such as finding the highest point and steepest slope on a path. However, the focus of the study was on determining the effect of time pressure on user interaction, and as such the research provides no information as to the relative usability of 3D and 2D depictions. In response to this deficit, Herman and Stachon (2016) compared the performance – measured in terms of efficiency and accuracy – of static 3D perspective and interactive 3D terrain depictions. The authors tasked 22 participants with solving a series of cartometric tasks (e.g. finding the highest point). Overall, static depictions were found to lead to significantly faster and more accurate user interactions.

However, methodological issues raise doubt about this conclusion. Herman and Stachon (2016) used a 'between-subject' approach, whereby participants were randomly split into two groups, each of which was exposed to one visualisation form (static 3D or interactive 3D). As such, the better performance of the static 3D display may be a function of the greater spatial ability of its group members. This is an especial shortcoming, given that Wilkening & Fabrikant (2013) and Cohen & Hegarty (2007) suggest that individual spatial ability can improve performance with 3D visualisations. Problematically, Herman & Stachon (2016) did not conduct any spatial ability tests on the study participants, meaning it is impossible to determine if this was indeed a factor in the result.

In light of the preceding literature, it is clear that the relative desirability of 2D and 3D cartographic visualisations remains an open question. In addition, the conclusions of the majority of the existing literature suggests that the potential benefits of 2D and 3D elevation visualisations are heavily dependent upon task use and context. As such, additional research into the relative usability of 2D and 3D elevation visualisations for cycle route-planners – the broad aim of this thesis defined in Section 1.2 – is therefore of prime importance. The following research objectives and questions presented in Chapter 4, and the associated methodology presented in Chapter 5, provide a research framework to address this aim.

4 Research Questions

In light of the preceding literature review, the following specific research questions have been developed, in order to help meet the broad aim of the study (Section 1.2). Given that the focus of this research project centres upon cartographic usability, the questions themselves are centred around ISO Standard 9241-11's components of usability – efficiency, effectiveness and user preference:

1. *Is there a significant difference in the relative efficiency of 2D and 3D elevation visualisations for cycle route-planners, when performing route-planning tasks related to:*

- *a. height detection?*
- *b. slope detection?*
- *c. overall climbing?*

2. *Is there a significant difference in the relative effectiveness of 2D and 3D elevation visualisations for cycle route-planners, when performing route-planning tasks related to:*

- *a. height detection?*
- *b. slope detection?*
- *c. overall climbing?*

3. *Do users prefer 2D or 3D elevation depictions when performing cycle route-planning tasks?*

The sub-questions (height detection, slope detection and overall climb estimation) were chosen based on the reviewed literature in Chapter 3. Determination of point height has been utilised as an empirical research question in a number of previous map-design research studies, including many of those reviewed in Section 3.5 (e.g. Phillips et al. 1975; Savage et al., 2004; Bleisch and Dykes, 2008; Popelka and Brychtova, 2013; Herman and Stachon, 2016). This user task, in addition to slope determination, was also a question included in the empirical study of Brügger et al (2016), which this thesis aims to directly build upon. The inclusion of similar question-types in this study should allow clear comparison with the results of this existing associated research.

However, one of the major criticisms of empirical map-design research is that the questions asked of participants are often not typical of real-life problems (Petchenik, 1983; Ciolkosz-Styl, 2012). Though research into cycling motivation showed that elevation and slope are indeed crucial determinants (Hochmair 2004; Su et al. 2010; Winters et al. 2011; Huffman 2009), asking users to assess the height or slope of single points in isolation is arguably an untypical cycle route-planning scenario. Surely, when a user is planning a cycling route, they are also interested in the route as a whole, in addition to any singular steep or high sections? Hence, the inclusion of the 'overall climb' research question in this thesis. By asking users to consider the whole route, it is hoped to gain a measure of visualisation-usability which is more representative of actual route-planning tasks.

5 Methodology

The research methodology was designed to respond specifically to the research questions posed in Chapter 4, whilst also addressing some of the methodological flaws apparent in the research reviewed in Chapter 3. Given the issues associated with expert-testing, and the greater suitability of empirical studies for assessing usability (Section 3.2), an empirical user-study was designed and conducted in March 2017, in order to address the study aim and research questions. The following section provides a broad contextual overview of the implementation of this user-study, before going on to provide in-depth methodological detail.

5.1 *Research Design*

5.1.1 *Broad Overview of Methodology*

In order to make the results comparable with the most closely associated existing research – the work of Brügger et al (2016) – a similar research design was implemented, albeit with some notable differences. Following the vast majority of previous cognitive map-design research (Chapter 3), a questionnaire-based survey was conducted. A series of interactive 2D and 3D map stimuli were firstly produced, using a range of different elevation visualisations – line width (2D), colour (2D), arrow (2D), elevation profile (2D), elevation profile (3D) and terrain (3D). Together, these formed the independent variables of the study. Users were then tasked with solving a series of route-planning questions about each of the visualisation types, in a controlled experimental environment. For each visualisation type, participants answered questions concerned with: height detection; slope detection; and overall climb determination.

The responses of the study participants – measured in terms of their accuracy (percentage of correct answers) and efficiency (response time) – formed the study's dependent variables, which were then used as the basis for addressing research questions 1 and 2. Before and after the survey, participants were also asked about their preferences for the visualisations in question; these responses were used to address research question 3.

5.1.2 *Experimental Design*

A repeated measures (within-subject) design was chosen as the experimental framework for the survey. As such, every participant answered questions about every one of the visualisation types. This framework was chosen over a between-subjects design, to ensure that the study results did not simply reflect individual differences in spatial ability, as was potentially an issue in the work of Herman & Stachon (2016). Instead, a repeated measures design should ensure that changes in user response are indicative of changes in the elevation visualisation type, thus providing direct insight into their relative usability.

As noted, the repeated-measures design implemented in Brügger et al. (2016) may have induced learning effects which compromised the results (Section 3.3). The potential cause was the use of a constant spatial area at the task level. For example, users were asked to select the highest of three points, on a map of Boston that used arrow visualisation. Later, they were asked again to select the highest of *the same three points*, but with a map that used an elevation profile. Clearly, there is potential for a learning effect in this instance. A solution here would be the use of different spatial areas for each visualisation type (Table 5-1). However, if there are greater terrain variations within one spatial area than another, it is feasible that changes in the recorded dependent variables (accuracy and efficiency) may be due to these spatial differences, rather than differences in visualisation usability.

To counter this issue, a group approach was taken in this study, whereby participants were randomly assigned to one of six groups. Within each group, a new city was used for each of the six visualisation types. For each group, the city and visualisation combination was changed. This design is summarised in Table 5-2. Using this design, any differences in

CITIES	TASKS		
	Highest Point *(which point is highest?)*	**Steepest Slope** *(which point lies on the steepest slope?)*	**Spatial cognition** *(which route has the overall highest height gain?)*
Bristol (Area a)	V1	V1	V1
Bristol (area B)	V2	V2	V2
Newcastle (area A)	V3	V3	V3
Newcastle (area B)	V4	V4	V4
Sheffield (area A)	V5	V5	V5
Sheffield (area B)	V6	V6	V6

V1 = Arrow visualisation, V2 = Line width visualisation, V3 = Colour visualisation, V4 = 2D elevation profile, V5 = 3D elevation profile, V6 = 3D terrain

Table 5-1 *A potential experimental design, whereby a different city is used for each visualisation.*

CITIES	TASKS		
	Highest Point *(which point is highest?)*	**Steepest Slope** *(which point lies on the steepest slope?)*	**Spatial cognition** *(which route has the overall highest height gain?)*
Bristol (area A)	A, 3T, 3P, 2D, C, L	A, 3T, 3P, 2D, C, L	A, 3T, 3P, 2D, C, L
Newcastle (area A)	L, A, 3T, 3P, 2D, C	L, A, 3T, 3P, 2D, C	L, A, 3T, 3P, 2D, C
Sheffield (area A)	C, L, A, 3T, 3P, 2D	C, L, A, 3T, 3P, 2D	C, L, A, 3T, 3P, 2D
Bristol (area B)	2D, C, L, A, 3T, 3P	2D, C, L, A, 3T, 3P	2D, C, L, A, 3T, 3P
Newcastle (area B)	3P, 2D, C, L, A, 3T	3P, 2D, C, L, A, 3T	3P, 2D, C, L, A, 3T
Sheffield (area B)	3T, 3P, 2D, C, L, A	3T, 3P, 2D, C, L, A	3T, 3P, 2D, C, L, A

✝ A = Arrow visualisation, L = Line width visualisation, C = Colour visualisation, 2D = P elevation profile, 3P = 3D elevation profile, 3T = 3D terrain

✝ Group 1, Group 2, Group 3, Group 4, Group 5, Group 6

Table 5-2 *The final 'group' experimental design, designed to avoid both learning effects and the influence of terrain differences between visualisation types.*

response accuracy and efficiency induced by a terrain differences rather than visualization differences, should be equally present in the data for all visualization forms, and thus overall should have negligible impact. In essence, the study therefore used a mixed design, featuring independent variables within subjects (visualisation type), and between subjects (spatial area). However, we are only truly interested in effects of the former.

5.1.3 *User-Study Questions*

The usability of each visualisation type was assessed using a series of targeted questions, which were asked for every visualisation form (as per Table 5-2). A criticism levelled at much cognitive map-design research is that the questions used during user-testing frequently fail to reflect the real life challenges that maps are used to answer (Petchenik 1983; Ciolkosz-Styk 2012). The questions used throughout this study attempted to avoid this pitfall.

1. 'Which of these points is highest?'

The goal of this question was to assess the relative usability of the designs for communicating height information (Figure 5-1). The same question was used in the study of Brügger (2015), who reasoned that cyclists would want to find the highest spot on a route, in order to take a panoramic photograph. Given that the focus therein and herein lies on cycling maps for urban areas, whereby such panoramic views are likely obscured by buildings, this reasoning is questionable. Nonetheless, research into cycling habits suggests that hills – and by association height gain – are key deterrents to cyclists (Hochmair 2004). Hence, determination of the height along a route is considered a typical route-finding question for a cyclist. For each visualisation, users were presented with three randomly selected points, each of which was located on one of three alternate routes (refer to Section 5.2.1.2 for further information on visualisation design). The points were randomly labelled with the letters

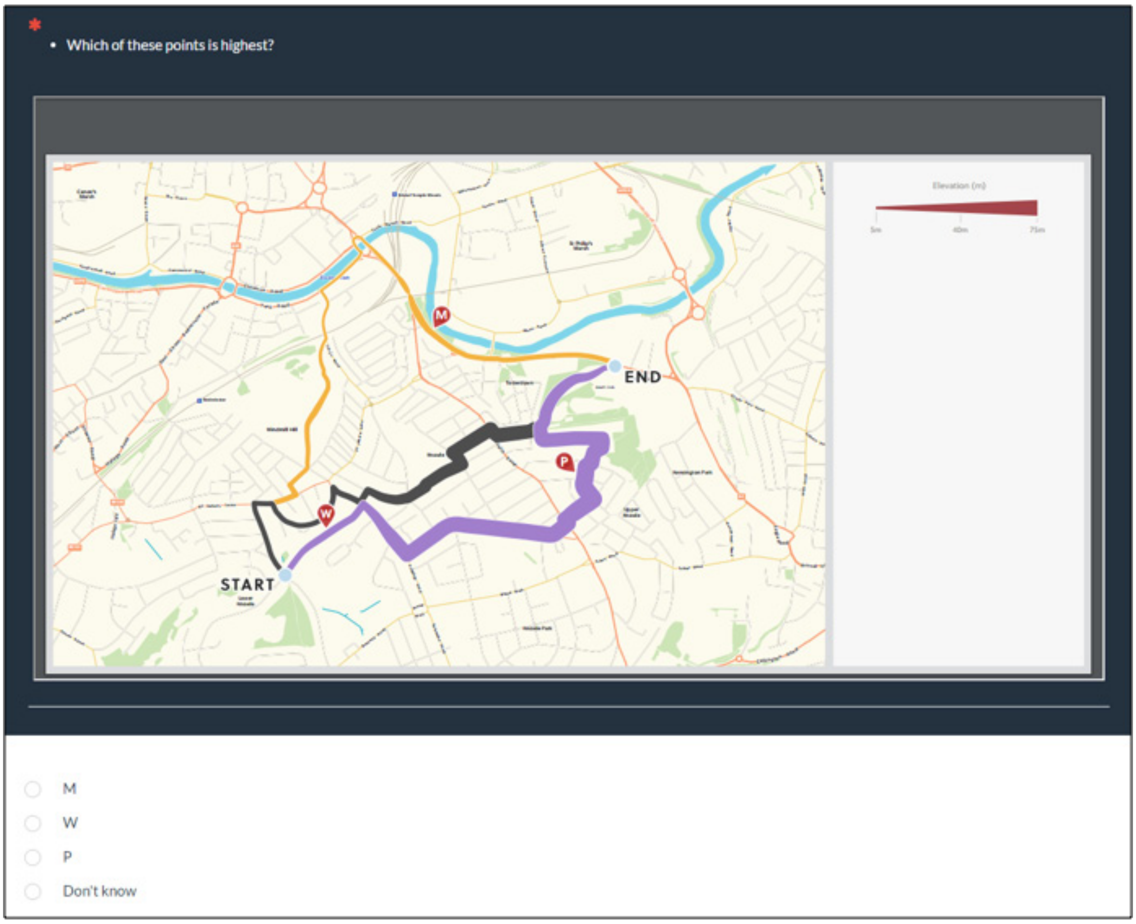


Figure 5-1 *An example of a 'point height' question, using line-width visualisation for three routes in the city of Bristol.*

D – Y, to avoid positive or negative associations, and were presented to test participants in a randomised order, to further negate learning effects. Users were then asked to determine which of the three points was the highest (Figure 5-1).

2. 'Which of these points lies on the steepest slope?'

The goal of this question was to assess the relative usability of the designs for communicating slope information. The steepness of a route is arguably of greater importance to the majority of cyclists than pure elevation (Hochmair 2004; Su et al. 2010; Huffman 2009). Indeed, Winters et al. (2011) found that flat routes were strongly influential in encouraging cyclists to make a journey. Thus, users were asked to select the point which was situated on the steepest slope, from three potential options. Again, the points were

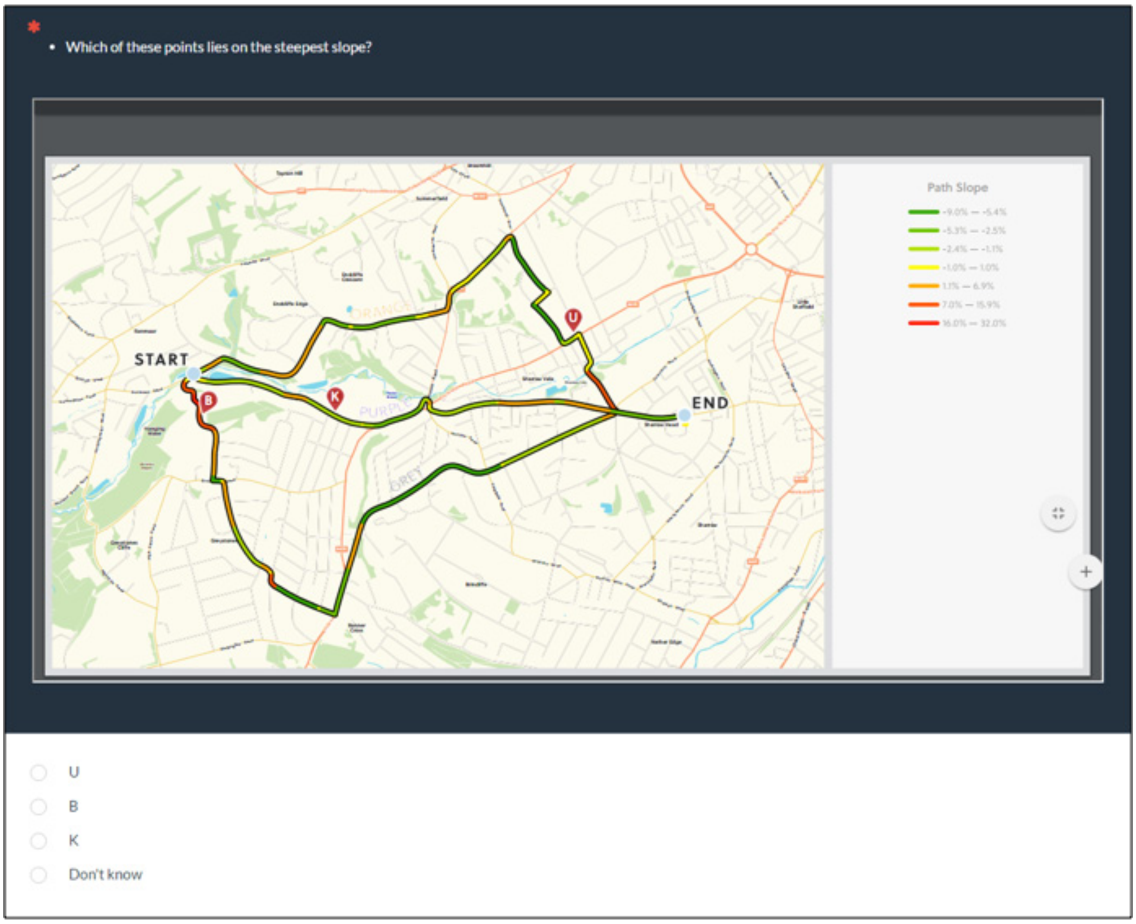


Figure 5-2 *An example of a 'point slope' question, using colour visualisation for three routes in the city of Sheffield.*

randomly situated, labelled, and presented, with one point located on each of three routes (Figure 5-2). The location of the slope points was kept distinct from the height points, to avoid any assumptions of equivalence by users.

3. 'Which of these routes includes the most amount of climbing?

The goal of this question was to assess the relative usability of the designs for communicating an overall impression of the spatial environment, and the terrain characteristics of entire routes. While the preceding questions asked participants about distinct map features, this question was intended as means of judging how well the visualisation was able to convey an overall impression of terrain to the user. For each question, users were shown three routes simultaneously, each of which had been randomly coloured, or labelled, 'orange',

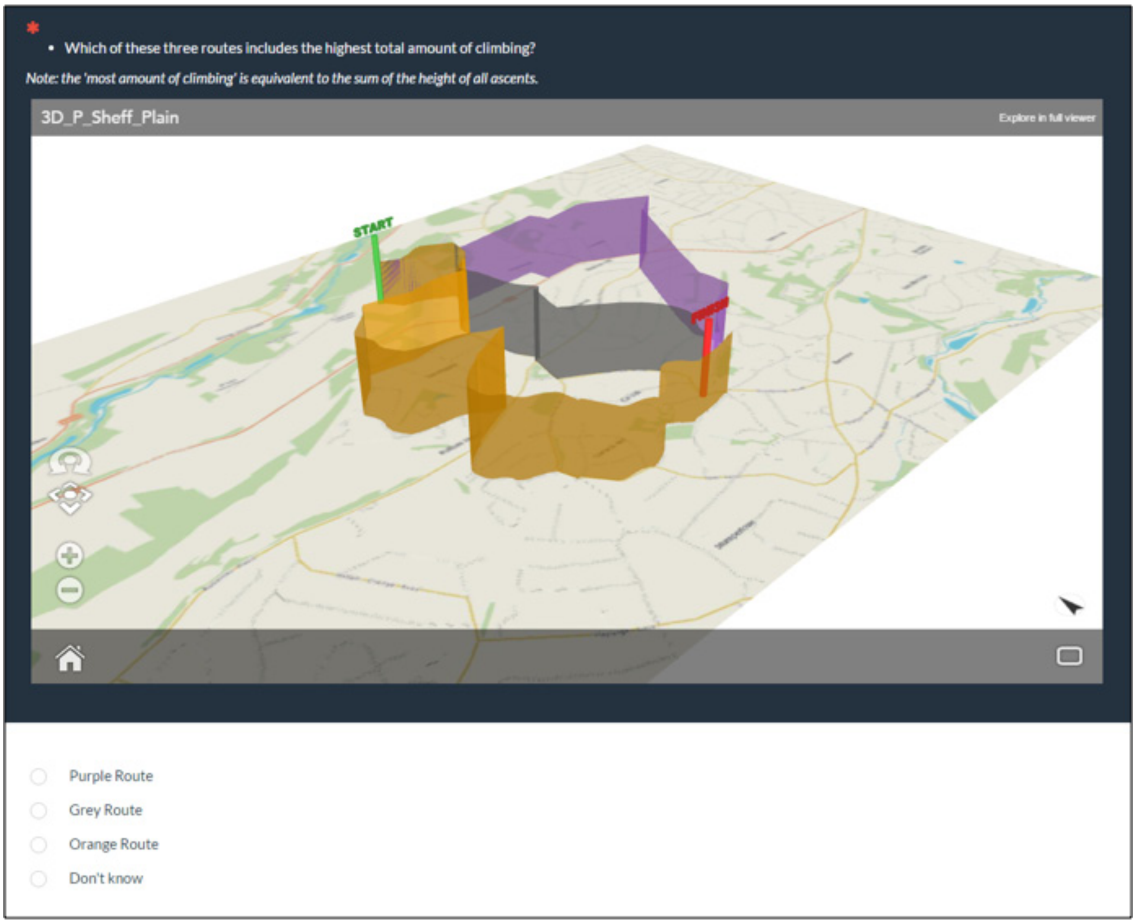


Figure 5-3 *An example of a 'climb' question, using a 3D elevation profile visualisation for three routes in the city of Sheffield.Bristol.*

'purple' and 'grey', and asked to select the route which contained the highest total amount of climbing. In pre-testing, this phrase 'most amount of climbing' caused some confusion; the clarification, "the 'most amount of climbing' is equivalent to the sum of the height of all ascents" was therefore added to the questionnaire prior to final user-testing (Figure 5-3).

5.2 *Materials*

5.2.1 *Map Stimuli*

All map stimuli used basemaps which covered geographic areas of a constant size (8.73km^2), and used in every case a constant scale of 1:10,000. In total, six geographic areas were chosen for the study: two each from the British cities of Sheffield, Newcastle and Bristol. These cities were chosen as they are fairly hilly, have good data coverage, and are unlikely to be familiar to the user-study participants in Austria. Each geographic area contained three alternative routes, all of which started together and ended together. This is characteristic of online route planners such as Google Maps (Section 3.4), and was the approach taken by Hochmair (2004) and Brügger (2015). Map stimuli were produced for every combination of spatial area, visualisaton type and question type (as summarised in Table 5-2), meaning that in total 108 separate visualisations were produced for the user-study.

5.2.1.1 *Data Sources*

Topographic map data for the production of the map-stimuli basemaps were obtained from the United Kingdom national mapping agency, The Ordnance Survey, via its 'OS Open Data' service (Ordnance Survey 2016). Detailed vector data (supplied as 'OS Open Map – Local') were downloaded to cover the extent of the three cities selected for the study, and were then processed using ArcMap 10.3. A custom symbolisation was then applied (removal of buildings, simplification of colour scheme, inclusion of cycle paths) to form appropriate basemaps that were typical of cycle route-planners (e.g. Figure 5-1). Throughout, the intention was to make the maps as 'typical' as possible, to avoid the issues of 'quasi-maps' discussed by Ciolkosz-Styk (2012), which was apparent in much early cognitive map-design research (Section 3.1). Hence, symbolisation and levels of abstraction were chosen to mimic the existing cycle route-planners reviewed in Section 3.4. These data were chosen as they were freely available to download under a UK Open Government Licence, and had complete coverage for the chosen geographic areas. As they were produced by a trustworthy national mapping agency, their accuracy and quality may also be assumed.

Airborne composite LIDAR terrain data (1m resolution) was obtained for all required geographic areas from the United Kingdom Environment agency (Environment Agency 2016). Again, data were selected as they were of high quality and accuracy, and available under an Open Government Licence. Coverage was also largely continuous and complete for the regions required, while 1m resolution was considered sufficient for the purposes of the study.

The data were obtained as raw ASCII files, and were mosaicked in an ArcGIS batch process to produce raster DTMs. Any holes in the data were filled using nearest neighbour sampling, to form a continuous surface. The raster DTM was then clipped to each geographic test area, and used as the basis for creating the various elevation visualisations (Section 5.2.1.2).

Route data for the paths displayed in each visualisation were created using the route-plotting service available on the online 'BikeHike' service (Bike Hike 2017). Routes were exported as gpx files, and imported into ArcGIS using the 'GPX to Features' tool. The (3D polyline) features that were created contained no elevation information; height data derived from the Raster DTMs were added to the route polylines using the ArcGIS 3D Analyst toolbox.

5.2.1.2 *Choice and Production of Elevation Visualisations*

Six different forms of elevation visualisation were tested, comprising a mix of 2D and 3D designs, identified both throughout the academic literature and in publicly accessible route planners:

a.	*Arrow visualisation*	*d.*	*2D elevation profiles*
b.	*Line-width visualisation*	*e.*	*3D elevation profiles*
c.	*Colour visualisation*	*f.*	*3D terrain models*

The elevation depictions were chosen primarily due to their existing popularity within online cycle route-planners (Section 3.4). In some instances, namely for line-width elevation visualisation and 3D route profiles, the depictions are not yet widely adopted, but nonetheless represent novel new solutions whose usability is worthy of testing. Further justification for the selection of visualisations is given in Section 3.4.

a. Arrow visualisation:

In this visualisation, arrows were used to indicate the presence of three grades of uphill slopes, as is typical of many cycling maps (Section 3.4). In general, slopes were classified as: slight uphills (1.1% – 5.0%); moderate uphills (5.1% – 10.1%); and steep uphills (10.2% – 30%), though the exact classification of the grades varied slightly across spatial areas. The path slope was derived in ArcMap, using the ET Surface Toolbox plugin (Tchoukanski 2016). This allowed each route polyline to be divided into segments based on path slope. The plugin splits polylines whenever the slope changes direction (i.e. from uphill to downhill, downhill to flat, and vice versa). The resulting uphill segments were then symbolised according to their average slope, using manual placement in Adobe Illustrator, and a legend added to the right of the map (Figure 5-4). Unlike in the work of Brügger (2015), arrows represented slope rather than segmented elevation gain. This decision was taken, as slope identification is the method overwhelmingly used in published cycling maps (Section 3.4).

Figure 5-4 *An example of an arrow-based terrain visualisation, in this instance for the city of Sheffield.*

Figure 5-5 *An example of a line-width terrain visualisation, in this instance for the city of Sheffield.*

b. Line-width visualisation:

This visualisation is based on the work of Huffman (2009), who designed a method for encoding elevation using line width (Figure 5-5). Using this symbology, higher elevations are encoded as larger line widths. Beneficially, this also means that slope can be inferred from the steepness of the edge of the line. Though the author characterises this symbology as a 'simple and effective' means of communicating elevation information (Huffman, 2009:86), no evidence to support this claim was provided, and the design has not yet undergone any user testing. As such, it was considered a worthy candidate for inclusion in this user study.

The line-width visualisations were produced using a combination of ArcGIS 10.3 and Adobe Illustrator CS5. ArcGIS was used firstly to create 10m contour lines from the DTM raster files for each city. The intersection between route polylines and these contour lines was then found, and symbolised with a circle proportional to the height of the route at that point (Figure 5-6a). Data were then exported to Adobe Illustrator, where the variable stroke width tool was used to match the width of the route path to the elevation at each intersection (Figure 5-6a). A legend was provided in the right hand section of the map for clarification.

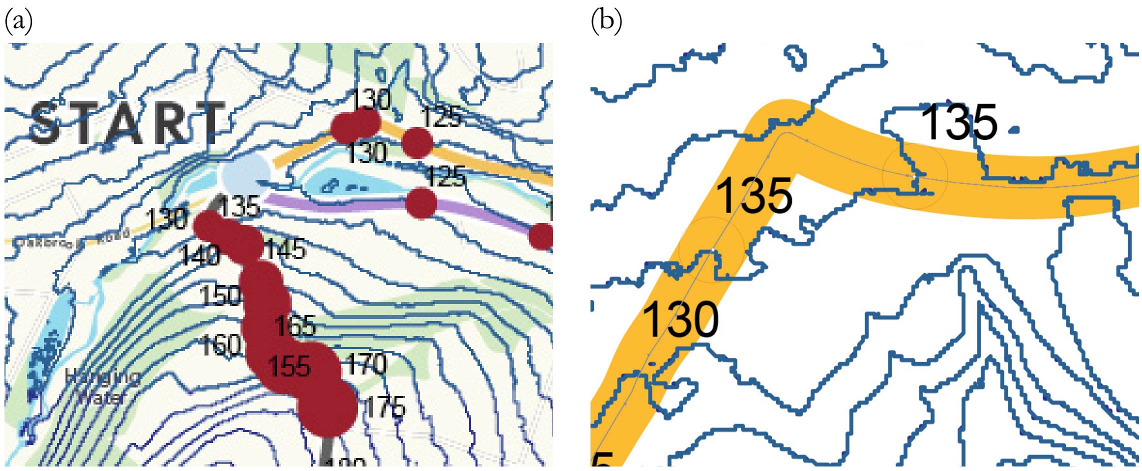


Figure 5-6 *Production of the line-width visualisation: (a) drawing circles proportional to height, at contour intersects; (b) adjusting line width to equal circle-diameter in Adobe Illustrator.*

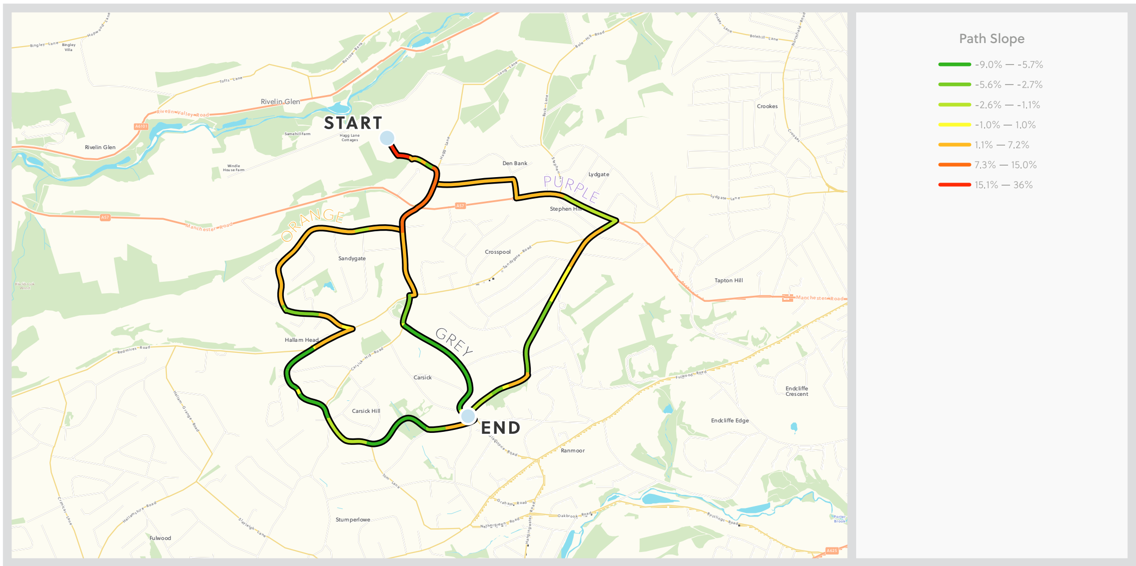


Figure 5-7 *An example of a colour visualisation, in this instance for the city of Sheffield.*

Figure 5-8 *An example of a 2D profile visualisation, in this instance for the city of Sheffield.*

c. Colour visualisation:

The visualisation was chosen due to its prominence in online route-planning software (Section 3.4). Production was initially similar in process to the arrow visualisation (c); the ET surface toolbox in ArcGIS was again used to segment lines on the basis of their varying slope. However, in this instance, all slopes (both uphill and downhill) were symbolised using a red-green colour spectrum, with a legend provided in the right-hand-side perimap region (Figure 5-7).

d. 2D elevation profiles:

2D elevation profiles, similar to those detailed in Section 3.4, were produced from the 3D polyline data for each route. Tabular height data were exported to Microsoft Excel and a rough profile produced, which was then styled in Adobe Illustrator. Separate profiles were produced for each route, and displayed in the right hand side perimap region. Profiles could be easily matched to their associated route by colour (Figure 5-8). Unlike the work of Brügger (2015), profiles and routes were not 'segmented' into height-based portions.

e. 3D elevation profiles:

The 3D elevation profile visualisation effectively consisted of the 2D profile, 'transformed' into a 3D depiction, and applied along the length of each route. This relatively new technique is slowly growing in popularity, but its usability is not yet understood (Section 3.4). The 3D profile was produced in ArcScene, then exported as a 3D Web Scene to permit hosting via ArcGIS online. Online hosting was crucial for this study, as it allowed the 3D visualisation to be embedded into the digital user-survey (Section 5.2.2.2). The visualisation was fully interactive, permitting zooming, panning, and rotating via intuitive tools, which were displayed in the lower-left of the visualisation (Figure 5-9).

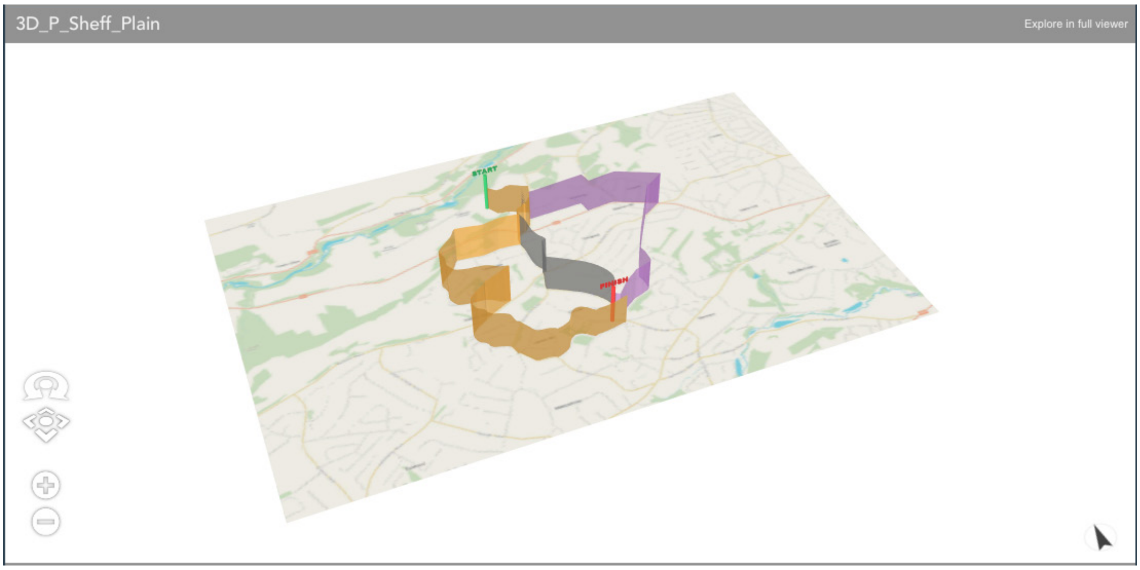


Figure 5-9 *A static screenshot of the interactive 3D elevation profiles depicting three routes from the city of Sheffield.*

Figure 5-10 *A static screenshot of the interactive 3D terrain model, and associated routes, for the city of Sheffield.*

f. 3D terrain models:

The 3D terrain model was produced in a similar fashion to the 3D elevation profile. Again, ArcScene was used for the layout, although in this instance the basemap was draped over a continuous DTM, forming a continuous terrain depiction. Each of the three routes 'floated' on this surface, and were visualised as a 3D tube. Though graphic continuous terrain depictions, such as contours, may be excessively occluded by other topographic features in urban areas, 3D terrain models do not suffer from the same issues. The model itself was illuminated under standard conditions (North-West azimuth, 45^{O} illumination altitude), and was hosted online in the same fashion as the 3D elevation profile (Figure 5-10).

5.2.2 *Apparatus*

5.2.2.1 *Hardware and Environment*

Wherever possible, standardised hardware and environmental conditions were used during testing, in order to minimise the impact of external environmental influences on the outcomes of the study. Thus, all user testing was performed on the same computer setup (Windows 10, 3.5GHz CPU, 16GB RAM, 3840 x 2160px display resolution on a Samsung 28" UE590 monitor), under constant artificial lighting conditions. All testing took place under supervised conditions, in a room within the Cartography Research Group, at the Technical University of Vienna.

5.2.2.2 *Software*

The user study was conducted using 'LimeSurvey' a free, open-source, online survey application. LimeSurvey is similar to SurveyMonkey, which has typically been used in many other cognitive map-design studies (e.g. Boer et al. 2012; Kalt 2015; Heim 2014). However, LimeSurvey offers a number of advantages. Unlike SurveyMonkey, it permits the inclusion of interactive HTML elements. As such, its use permitted the inclusion of the interactive 3D visualisations. Question randomization, to reduce learning effects, was also possible, as was automatic recording of user responses and response times.

The questions were presented in blocks (height, slope, climb, and user preference), each of which was preceded by a short introductory slide. Within the question blocks, individual questions were presented in a random order. Prior to the onset of the main questionnaire, users were also asked general background questions about their cartographic, geographic and educational background. They were also given the opportunity to familiarise themselves with all of the visualisation forms, and their interactive controls. In total, the survey consisted of 32 questions: 6 point height questions, 6 slope questions, 6 climb questions, 4 preference questions, and 10 background-information questions.

Pre-testing of the survey highlighted some design issues, which were resolved before the final survey was implemented. The 3D visualisations, hosted through ArcGIS online, took an extended and variable quantity of time to load (between 30 and 50 seconds). As such, the response time recordings for these questions were inaccurate, by an unknown value between 30-50 seconds. This issue was solved by providing an interim loading screen of a constant duration (60 seconds) to initially 'hide' the question, which was latterly removed to show the 3D visualisation. During analysis, the 60-second loading time was subtracted from the recorded response time, to give the 'true' response time.

5.3 *Study Procedure*

Every participant was booked into a 40 minute time-slot, which gave ample time for the entire experimental procedure. Participants were assigned randomly to each of the six question-groups (Section 5.1.2), such that each group contained an equal number of respondents. Prior to the participant's arrival, a computer was pre-loaded with the appropriate questionnaire. Participants were then greeted, and a general introduction to the survey was given, whereby they were informed of the following:

- The purpose and broad aim of the study.
- That their data would be anonymised.
- That their answers and response times would be recorded.
- The necessity to answer at a 'natural' pace, and to complete the survey in a single sitting without questions, pauses or interruptions.

Each questionnaire was designed to be as self explanatory as possible. Thus, after the introduction, participants were allowed to complete the survey independently, albeit under supervision. An example of one of the six questionnaires is included in its entirety in Appendix 1.

5.4 *Participants*

A total of 36 participants took place in the user study over a two-week period, stretching from the 6th to the 20th of March 2017. The majority of participants were recruited verbally, with the remainder booked through e-mail and via text-message. Around half were students on the International Masters programme in Cartography, with additional participants either personal friends of the author, or employed within the Cartography Research Group at the Technical University of Vienna. Following completion of the survey, participants were thanked for their participation with Toffifee caramel. A detailed profiling of the participants' background is given in the results section (Chapter 6).

5.5 *Methods of Statistical Data Analysis*

Data analysis was conducted using SPSS v.23, with a significance level of 5% ($p<0.05$) used throughout all statistical testing. Statistical tables were produced in Microsoft Excel 2011, and data visualisations were produced in SPSS v.23, then stylized in Adobe Illustrator CS5 where necessary. The presence of significant differences in user response time between visualisation types (Research Question 1) was tested using one-way ANOVA, firstly for data pooled for all tasks, and latterly task-by-task. Where a significant interaction was identified, Tukey's post-hoc testing was conducted to determine between exactly which visualisations significant differences were apparent. Prior to this statistical analysis, all response-time data were log-transformed in order to meet the conditions of normality (Shapiro-Wilk test) and homoscedacity (Levene's test) demanded by one-way ANOVA (Field 2009). One-way repeated-measures ANOVA was not used as, whilst the same participants were present in several groups (visualisation groups), they were tested in each instance on different spatial areas (Section 5.1.2), and thus the samples were not 'paired' as such in a repeated-measures design. A mixed-design ANOVA was not used as the influence of the between subjects factor (spatial area) was not of interest in this study.

Data concerning the accuracy of the different visualisation types (Research Question 2) were nominal ('correct' or 'incorrect' answers), and thus a chi-squared test of independence was used to determine if there was a significant interaction between visualisation type and result accuracy. Where a significant interaction was identified, multiple pairwise Fisher's-exact tests were conducted to determine between which visualisations types there was a significant difference. This test was used rather than chi-squared as the number of expected samples (that is, total correct or incorrect answers for a visualisation type) was in some cases less than 5 (McDonald 2014). Again, this analysis was conducted firstly for data pooled from all tasks, and then again on a task-by-task basis.

6 Results

The following section aims to provide an objective and descriptive presentation of the results of the study. Beginning with information on the educational, knowledge and demographic backgrounds of the participants, the section subsequently provides detailed results on the relative efficiency (response time), effectiveness (accuracy) and user preference of the visualisations detailed previously.

6.1 *Participants – Background Information*

The 36 participants who took part in the survey were comprised of a non-uniform mix of genders (25 females and 11 males) and ages, with most aged between 21–30 (Appendix 2). No participants suffered from any visual impairments, aside from vision corrected using eyeglasses, and none had previously visited any of the three cities which formed the basis of the test-stimuli. All the participants were well educated, with all-bar-one having already obtained a graduate or postgraduate degree. Given this context, it is unsurprising that the vast majority of participants had either high or expert knowledge of reading paper maps (75% of participants) and digital maps (83.3% of participants). Similarly, a fairly large proportion (61.1%) of respondents considered themselves either highly trained or experts in Cartography and GIS (Figure 6-1). However, of note is that participant knowledge of both 3D maps and 3D graphics in general is comparatively poor.

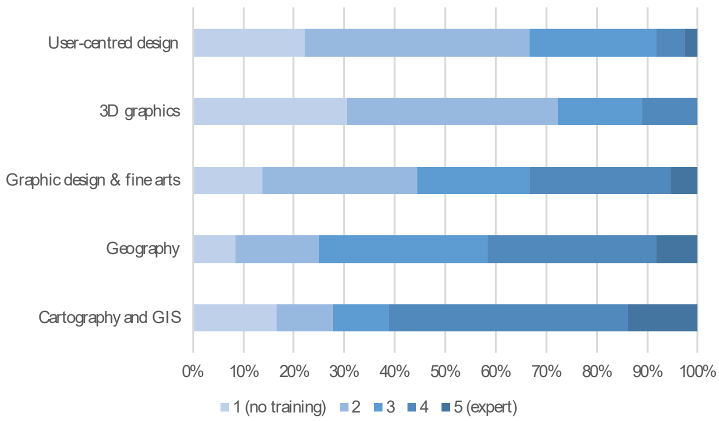


Figure 6-1 *A summary of the (self-assessed) knowledge and experience levels of the participants, in fields related to the study and user-test.*

6.2 Efficiency (response time)

6.2.1 Combined Tasks

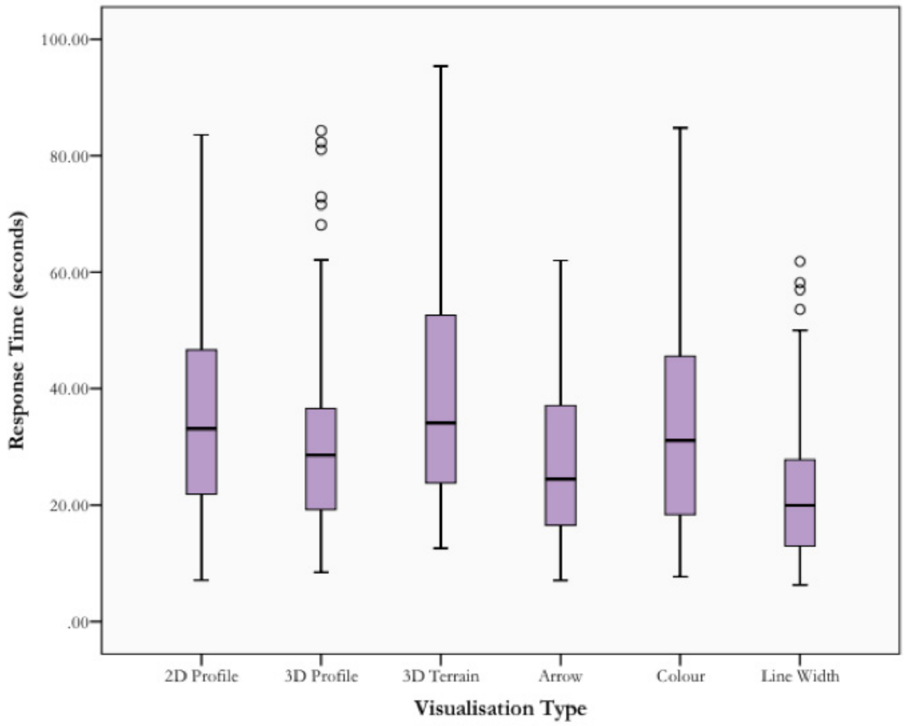


Figure 6-2 *Box-whisker plot showing the central tendency and spread of response times for each visualisation type (data from all tasks – height, slope and climb – combined). Small circles represent data outliers.*

Data for all tasks were firstly pooled, and analysed together. Initial visualisation of these data (Figure 6-2) suggested that user response time varies along with visualisation type. Subsequent one-way ANOVA confirmed that there was a significant difference in user response time between different visualisation types ($F(5,642) = 15.9$, $p = .000$). Tukey's post-hoc testing was then conducted to determine between which visualisation types there was a significant difference in response time (Table 6-1).

	Combined (all tasks)					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		0.700	0.435	**0.010**	0.648	**0.000**
3D Profile			**0.013**	0.377	1.000	**0.000**
3D Terrain				**0.000**	**0.010**	**0.000**
Arrow					0.427	**0.040**
Colour						**0.000**

Table 6-1 *Results of Tukey's post-hoc analysis for data pooled from all tasks, identifying between which visualisation types there was a significant difference in response time. Statistically significant differences ($p<0.05$) are highlighted in bold.*

User response time with line width visualisation (M = 22.5, S.E. = 1.15) was significantly lower than for all other tested forms of elevation visualisation, both 2D and 3D. The next most efficient visualisation type was arrow visualisation (M = 28.0, S.E. = 1.40), whose response time was significantly lower than the 2D profile (M = 35.7, S.E. = 1.7), 3D terrain (M = 40.3, S.E. = 2.00) and colour (M = 32.8, S.E. = 1.73) visualisations. Comparing 3D representations, the response time for the 3D terrain visualisation was significantly higher than for 3D profile visualisations (M = 31.5, S.E. = 1.54). Indeed, response time with the 3D terrain visualisation was significantly higher than for all other tested visualisations, aside from the 2D profile. While the mean response time for 3D Terrain was also slightly higher than the 2D profile, this difference was not statistically significant.

6.2.2 Height Tasks

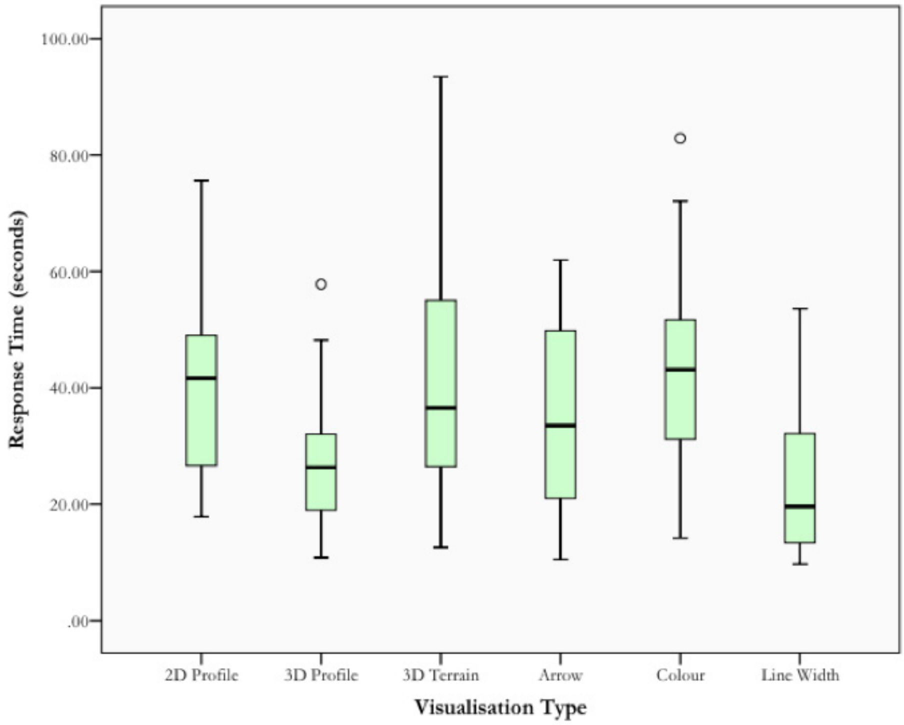


Figure 6-3 *Box-whisker plot showing the central tendency and spread of response times for each visualisation type, for 'height' tasks only. Small circles represent data outliers.*

When response-time data were analysed at a task-by-task level, there were some notable differences in the outcomes, relative to the pooled data. For height tasks, box-whisker plots again suggested differences in response time between visualisation types (Figure 6-3).

The significance of these differences was confirmed with a one-way ANOVA ($F(5,210) = 11.8$, $p = .000$). Post-hoc testing (Tukey's) was then conducted to elucidate the statistical significance of any between-visualisation differences (Table 6-2).

	Height Task					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		0.060	1.000	0.561	1.000	**0.000**
3D Profile			0.080	0.389	**0.002**	0.248
3D Terrain				0.625	0.999	**0.000**
Arrow					0.364	**0.001**
Colour						**0.000**

Table 6-2 *Results of Tukey's post-hoc analysis for height questions, identifying between which visualisation types there was a significant difference in response time. Statistically significant differences ($p<0.05$) are highlighted in bold.*

Again, pairwise comparisons demonstrate that height-task questions answered with the line-width visualisation generally had the lowest response time (M = 22.7, S.E. = 1.9). The response time for this visualisation type was significantly lower than that for the 2D profile (M = 40.6, S.E. = 2.72), 3D terrain (M = 42.2, S.E. = 3.59), arrow (M = 35.1, S.E. = 2.62), and colour (M = 42.2, S.E. = 2.75) visualisations. The mean response time was also slightly lower than that for the 3D profile visualisation (M = 27.5, S.E. = 1.76), although the difference was not enough to be statistically significant. Although the response time for the 3D terrain visualisation was again higher than the response time for the other tested 3D visualisation – the 3D elevation profile –the difference was not statistically significant. Likewise, there was no significant difference in the response times between 3D terrain, arrow and colour visualisations. Response time for the 3D profile, was significantly lower than that for the colour visualisation, but did not differ significantly from any other visualisation type, including line-width.

6.2.3 *Slope Tasks*

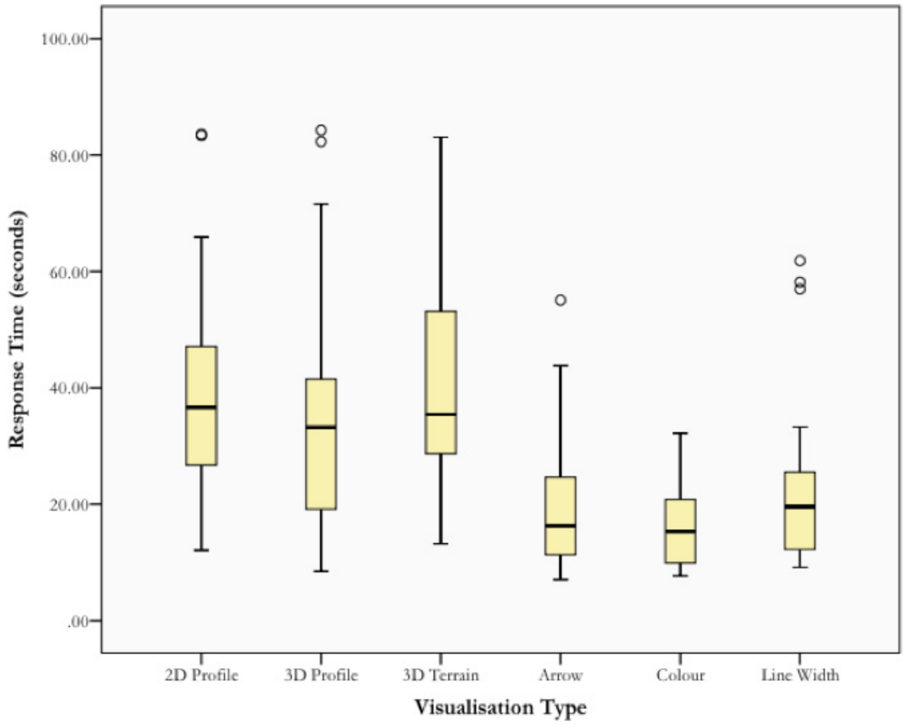


Figure 6-4 *Box-whisker plot showing the central tendency and spread of response times for each visualisation type, for 'slope' tasks only. Small circles represent data outliers.*

Visualisation of the response time data for slope identification tasks suggested differences between response time and visualisation type (Figure 6-4). This was confirmed with one-way ANOVA testing, which highlighted a significant interaction between visualisation type and response time ($F(5,210) = 22.3$, $p = .000$). Tukey's post-hoc testing (Table 6-3) gave further insight into the exact nature of the interactions.

	Slope Task					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		0.862	0.998	**0.000**	**0.000**	**0.000**
3D Profile			0.609	**0.000**	**0.000**	**0.000**
3D Terrain				**0.000**	**0.000**	**0.000**
Arrow					0.849	0.993
Colour						0.504

Table 6-3 *Results of Tukey's post-hoc analysis for 'slope' questions, identifying between which visualisation types there was a significant difference in response time. Statistically significant differences (p<0.05) are highlighted in bold.*

Broadly speaking, the response times for the 2D profile (M = 38.7, S.E. = 2.77), 3D profile (M = 35.5, S.E. = 3.22), and 3D terrain (M = 42.0, S.E. = 3.5), were much higher than for the arrow (M = 20.4, S.E. = 2.0), colour(M = 16.7, S.E. = 1.2), and line width (M = 21.6, S.E. = 2.2) visualisations. The results of Tukey's post-hoc testing revealed that these differences were indeed statistically significant. Hence, for slope identification questions, it is possible to group visualisation types into those with slower response times (2D profile, 3D profile and 3D terrain) and faster response times (arrow, colour, and line-width). Within these groups, there were no significant differences in response time.

6.2.4 *Climb Tasks*

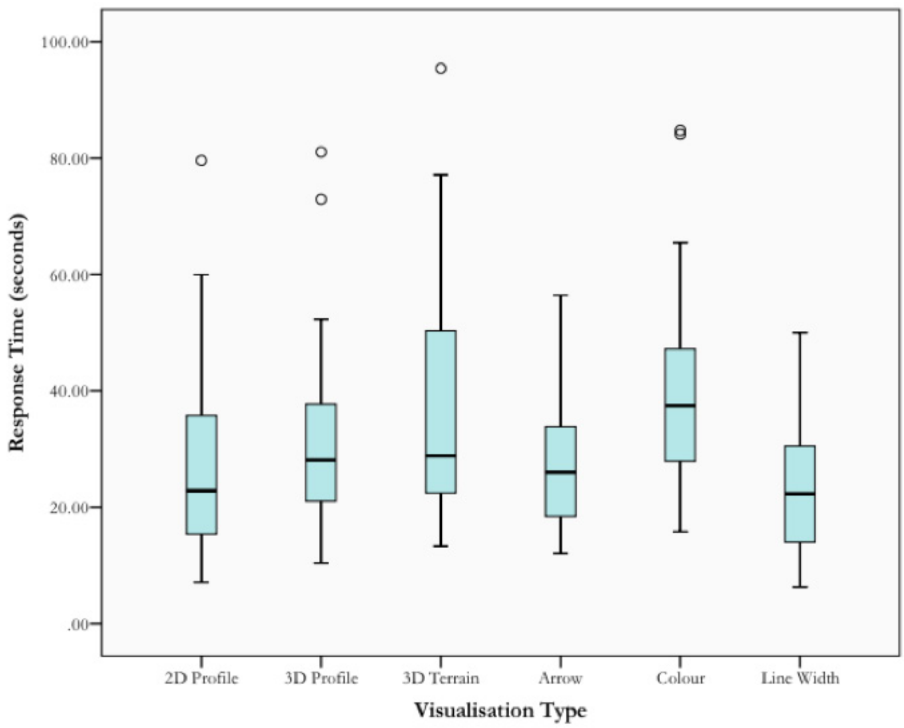


Figure 6-5 *Box-whisker plot showing the central tendency and spread of response times for each visualisation type, for 'climb' tasks only. Small circles represent data outliers.*

Visual analysis of the results (Figure 6-5) once again indicated that there were differences in response time between visualisation types, when answering 'climb' questions. Latterly, the results of a one-way ANOVA confirmed that some of these differences were statistically significant ($F(5,210) = 22.3$, $p = .000$). The results of Tukey's post-hoc testing (Table 6-4) highlighted those group-by-group pairwise interactions which were significant.

	Climb Task					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		0.532	**0.034**	0.844	**0.001**	0.908
3D Profile			0.780	0.995	0.217	0.073
3D Terrain				0.451	0.934	**0.001**
Arrow					0.066	0.235
Colour						**0.000**

Table 6-4 *Results of Tukey's post-hoc analysis for 'climb' questions, identifying between which visualisation types there was a significant difference in response time. Statistically significant differences ($p<0.05$) are highlighted in bold.*

The visualisation type with the lowest average response time was again line-width (M = 23.1, S.E. = 1.94), although response times with this visualisation type were not significantly different from questions answered with the 2D profile (M = 27.7, S.E. = 2.93), 3D profile (M = 31.4, S.E. = 2.72), or arrow (M = 28.5, S.E. = 2.03), visualisations. They were, however, significantly lower than 'climb' questions answered using 3D terrain (M =36.7, S.E. = 3.26) and colour (M = 39.5, S.E. = 2.72) visualisations. Both these visualisation forms – 3D terrain and colour – resulted in the longest response times when answering climb questions. Response times with both of these visualisation types were higher than from all other depictions, and significantly higher than those from 2D profile and line-width visualisations.

6.3 *Accuracy (effectiveness)*

6.3.1 *Combined Tasks*

A chi-square test of independence was conducted to ascertain if there was a relationship between visualisation type and the proportion of correct answers, for data pooled from all tasks. The results of this test highlighted a statistically significant interaction between these two variables (X^2 (5, N = 648) = 20.5, p = .001). Multiple pairwise comparisons using Fisher's exact test were then conducted to determine visualisation-by-visualisation relationships; these results are summarised in Table 6-5.

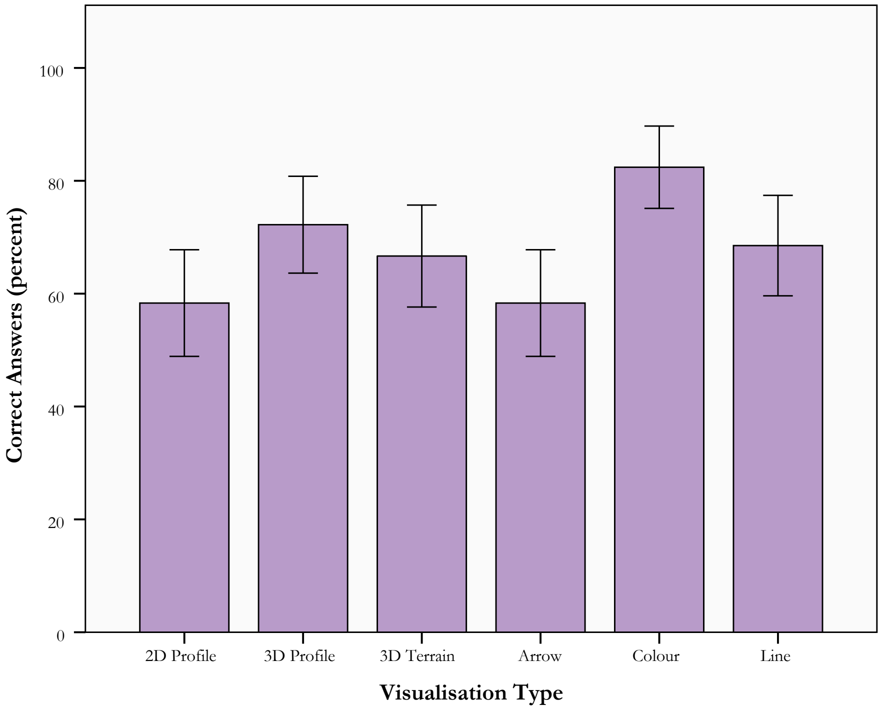


Figure 6-6 *The percentage of correct responses from the user study for each visualisation type, using data pooled from all question types ('height', 'slope' and 'climb'). Error bars represent 95% confidence intervals.*

	Combined (all tasks)					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		**0.045**	0.261	1.000	**0.000**	0.158
3D Profile			0.460	**0.045**	0.104	0.655
3D Terrain				0.261	**0.012**	0.884
Arrow					**0.000**	0.120
Colour						**0.026**

Table 6-5 *The results of multiple pairwise Fisher's exact tests, identifying between which visualisation types there was a significant difference in the proportion of correct responses, for data pooled from all questions. Statistically significant differences ($p<0.05$) are highlighted in bold.*

For all tasks combined, the colour visualisation resulted in a significantly higher proportion of correct answers (82.4%, S.E. = 3.66%) than all other visualisation types, aside from the 3D profile (72.2%, S.E. = 4.31%). Conversely, the proportion of correct answers for the arrow visualisation and 2D profile visualisations (58.3%, S.E. = 4.74%) was lower than for all other visualisation types, and significantly lower than for the 2D profile and colour visualisations. No other major patterns were apparent in the accuracy data pooled from all question types: there was no significant difference in the proportion of correct answers between the line (68.5%, S.E. = 4.47%), arrow (58.3%, S.E. = 4.74%) and 3D terrain (66.7%, S.E. = 4.54%) visualisations.

6.3.2 *Height Tasks*

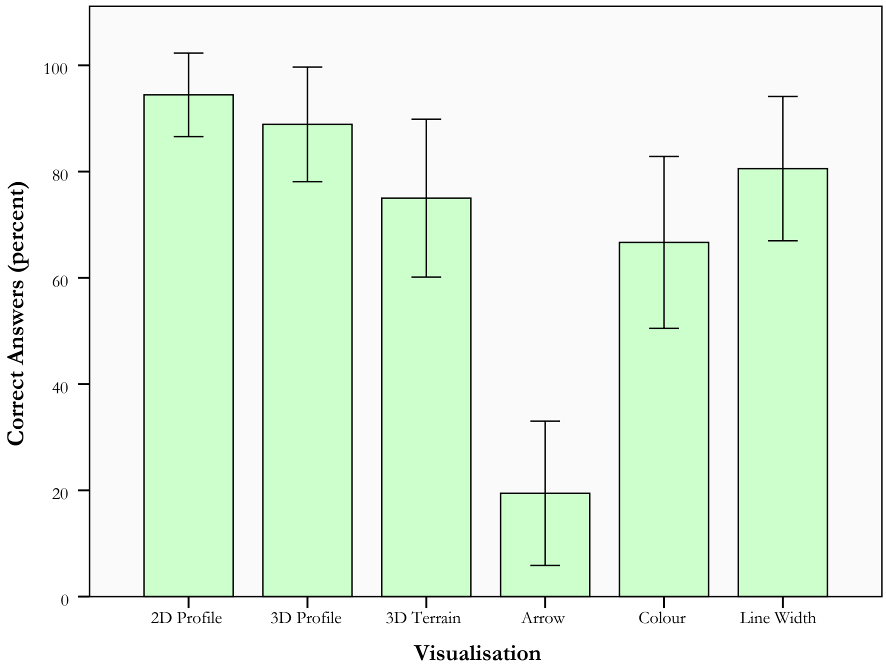


Figure 6-7 *The percentage of correct responses from the user study for each visualisation type, for 'height' tasks only. Error bars represent 95% confidence intervals.*

Accuracy data only from 'height' questions revealed a markedly different pattern of results (Figure 6-7). A chi-squared test for independence, using data only from 'height' questions, again revealed a significant association between the type of visualisation used, and the number of correct answers (X^2 (5, N = 216) = 61.1, p = .000). Unsurprisingly, pairwise comparisons using Fisher's exact test (Table 6-6) confirmed that the proportion of correct answers with the arrow visualisation (19.4%, S.E. = 3.81%) was significantly lower than for height questions answered with all other visualisation types. Though not as bad as the arrow visualisation, the colour visualisation also performed comparatively poorly with this task: its proportion of correct answers (66.7%, S.E. = 4.54%) was lower than all visualisation types aside from the arrow visualisation, and significantly lower than the 2D and 3D profiles (94.4%, S.E. = 2.20%, and 88.9%, S.E. = 3.02% respectively). Indeed, the large proportion of correct responses with the 2D profile was significantly higher than for all other visualisation types, aside from the 3D profile. No other notable trends were apparent in the data.

	Height Task					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		0.674	**0.046**	**0.000**	**0.006**	**0.046**
3D Profile			0.220	**0.000**	**0.045**	0.220
3D Terrain				**0.000**	0.605	1.000
Arrow					**0.000**	**0.000**
Colour						0.605

Table 6-6 *The results of multiple pairwise Fisher's exact tests, identifying between which visualisation types there was a significant difference in the proportion of correct responses, for data from 'height' questions only. Statistically significant differences (p<0.05) are highlighted in bold.*

6.3.3 *Slope Tasks*

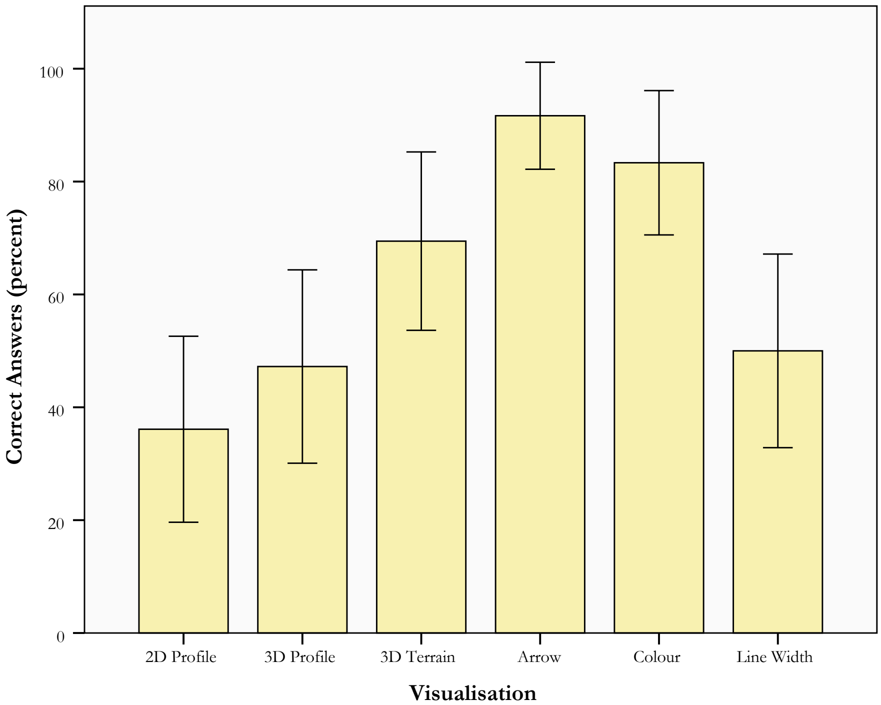


Figure 6-8 *The percentage of correct responses from the user study for each visualisation type, for 'slope' tasks only. Error bars represent 95% confidence intervals.*

A visualisation of the data representing the accuracy of 'slope' question answers only (Figure 6-8), appears to show an inverse pattern to that for height questions. Statistical analysis of this data (chi-squared test for independence) highlighted a significant interaction between visualisation type, and the number of correct answers (X^2 (5, N = 216) = 37.3, p = .000). Multiple pairwise comparisons using Fisher's exact test (summarised in Table Table 6-7) reveal the exact nature of visualisation-by-visualisation comparisons. The highest proportion of correct answers was for the arrow visualisation (91.7%, S.E. = 2.66%), which was significantly larger than the result for both 3D visualisation types (3D profile = 47.2%, S.E. = 4.80%, 3D terrain = 69.4%, S.E. = 4.43%), and also the 2D profile, which had the lowest accuracy for slope tasks (36.1%, S.E. = 4.62%). The colour visualisation also resulted in a high proportion of correct answers for slope questions (83.3%, S.E. = 3.59%). While this a higher proportion than for all other visualisation forms except the arrow visualisation, the difference was only statistically significant between the 2D profile and 3D profile visualisations (Table 6-7).

	Slope Task					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		0.474	**0.009**	**0.000**	**0.000**	0.341
3D Profile			0.094	**0.000**	**0.003**	1.000
3D Terrain				**0.035**	0.267	0.149
Arrow					0.478	0.000
Colour						0.478

Table 6-7 *The results of multiple pairwise Fisher's exact tests, identifying between which visualisation types there was a significant difference in the proportion of correct responses, for data from 'slope' questions only. Statistically significant differences ($p<0.05$) are highlighted in bold.*

6.3.4 *Climb Tasks*

For questions where users had to determine the total amount of climbing on different routes, the proportion of correct answers again varied alongside visualisation type (Figure 6-9). A chi-square test of independence uncovered a statistically significant interaction between visualisation type and the proportion of correct answers (X^2 (5, N = 216) = 30.1, p = .000). Subsequent pairwise comparisons using Fisher's exact test showed the exact nature of visualisation-by-visualisation comparisons, and where significant differences in accuracy occurred (Table 6-8).

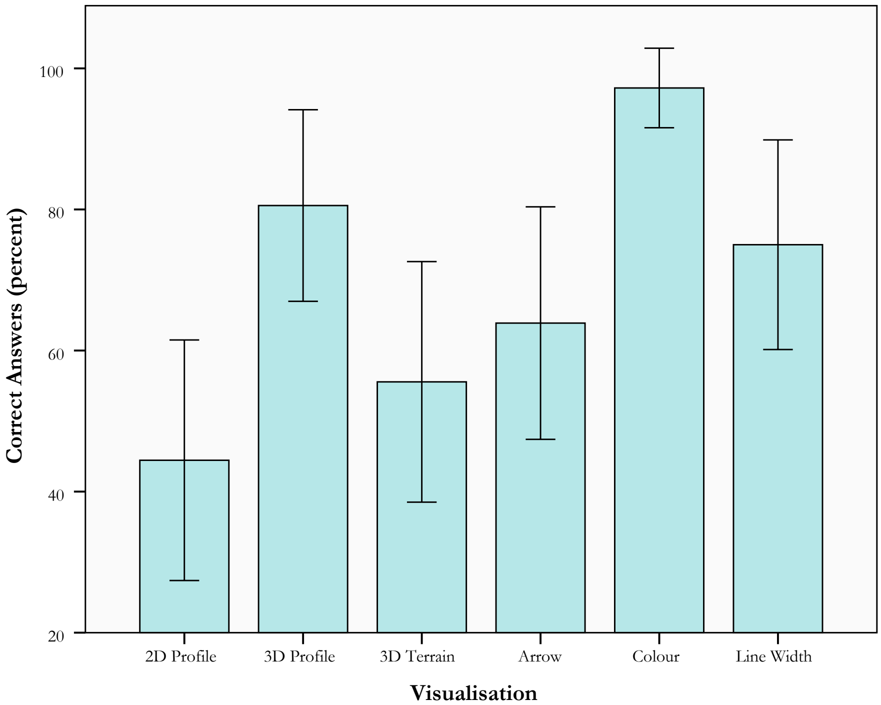


Figure 6-9 *The percentage of correct responses from the user study for each visualisation type, for 'climb' tasks only. Error bars represent 95% confidence intervals.*

	Climb Task					
	2D Profile	3D Profile	3D Terrain	Arrow	Colour	Line Width
2D Profile		**0.003**	0.480	0.155	**0.000**	**0.016**
3D Profile			**0.042**	0.188	0.055	0.778
3D Terrain				0.631	**0.000**	0.137
Arrow					**0.001**	0.443
Colour						**0.014**

Table 6-8 *The results of multiple pairwise Fisher's exact tests, identifying between which visualisation types there was a significant difference in the proportion of correct responses, for data from 'climb' questions only. Statistically significant differences ($p<0.05$) are highlighted in bold.*

The colour visualisation performed extremely well for this question type: 97% of questions answered using this visualisation form were correct (S.E. = 1.58%). This result was significantly higher than for all other visualisation types, except that for the 3D profile (80.6%, S.E. = 3.81%). By contrast, the proportion of correct answers with the alternative 3D visualisation – 3D terrain – was significantly lower (55.6%, S.E. = 4.78%). Indeed, only the 2D profile resulted in a lower proportion of correct answers (44.4%, S.E. = 4.78%). The result for the line width visualisation (75.0%, S.E. = 4.17%) significantly different only from the most accurate (colour) and the least accurate (2D profile) visualisations.

6.4 *User Preference*

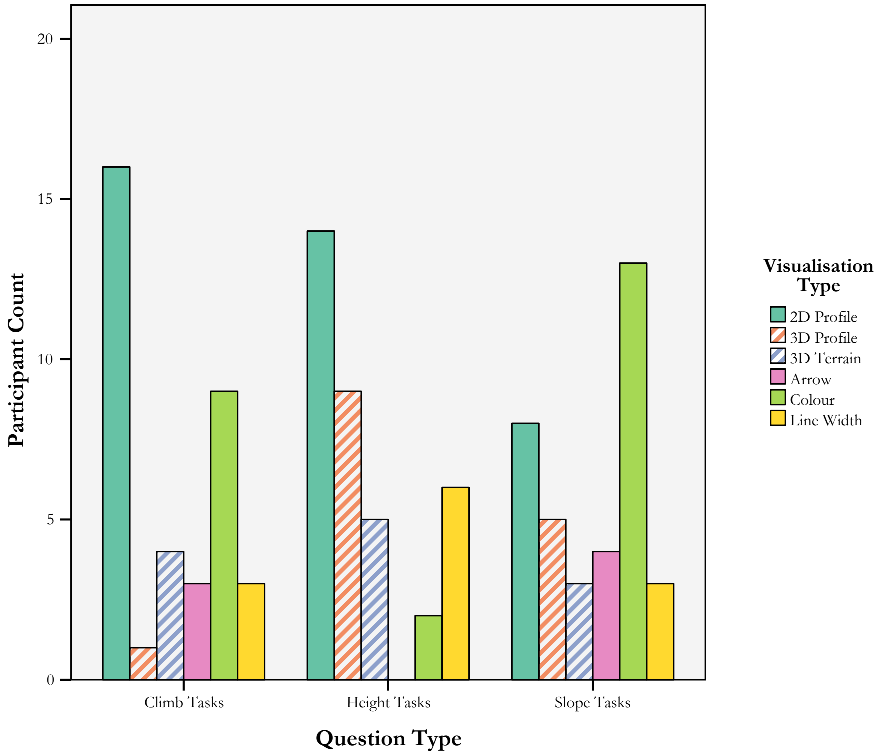


Figure 6-10 *A clustered bar-chart depicting visualisation preferences for the different question types assessed in the study. 3D visualisations are highlighted with a diagonal fill, and 2D visualisations in a solid colour.*

Following the assessment questions for height, slope and climb tasks, study participants were asked to state their preferred form of visualisation, for each of the different question types. The results are summarised in Figure 6-10. Beyond a simple choice of favoured visualisation type, participants were also asked to give reasoning for their choices – this information is discussed further in Chapter 7, whilst the full responses are included in their entirety in Appendix 3.

For all question types, the most popular visualisation type was a form of 2D rather than 3D visualisation. This trend was most apparent for the climb questions (only 2.8% of respondents (1 person) preferred the 3D profile, and 11.1% (4 people) the 3D terrain visualisation), and to a lesser extent for the slope tasks (13.9% preferred the 3D profile, 8.3% the 3D terrain visualisation). By contrast, for height questions, the 3D profile was the second most popular visualisation type (25% of respondents preferred this visualisation type). Notably, for height tasks, no participant chose the arrow visualisation as their favoured form of terrain depiction. A preference for arrow visualisations was also low for climb tasks (8.3%) and slope tasks (11.1%). The line-width visualisation was similarly unpopular for all question types.

At the opposite end of the preference spectrum, the 2D elevation profile was consistently popular for all question types. The 2D elevation profile was on average the favoured visualisation type for both climb tasks (44.4% of respondents) and height tasks (38.9% of respondents). Only for the slope tasks was the 2D profile not the most popular visualisation type, although it was second most popular. For this question type, 36.1% of participants preferred the colour visualisation. The colour visualisation was also popular for climb tasks, where it was the favoured choice of 25% of study participants.

7 Discussion

The following chapter discusses the results of the study in relation to the research questions posed in Chapter 4, and more broadly with reference to related literature discussed in Chapter 3. Latterly, additional findings outwith the remit of the thesis research questions are presented, and the wider implications of the study are discussed.

7.1 *Discussion in Relation to Research Questions*

7.1.1 *Research Question 1*

Q. *Is there a significant difference in the relative efficiency of 2D and 3D elevation visualisations for cycle route-planners, when performing route-planning tasks related to: (a) height detection; (b) slope detection; (c) overall route climb?*

The results of the study highlighted significant differences in efficiency – measured in terms of response time – not only broadly between 2D and 3D visualisations as a whole, but also at a smaller scale, between different forms of 2D and 3D depictions. The exact nature of these differences varied along with the nature of the route-planning task through which efficiency was assessed.

For questions of height detection, users answered most efficiently with a 2D visualisation – line width. Response time with this depiction was significantly lower than for all other visualisations, aside from the 3D profile, which also had a low response time. Thus, of the two most efficient visualisations for detecting height, one was 2D and one was 3D. By contrast, the (2D) colour visualisation was significantly less efficient than both these depictions. The slow response time for the colour visualisation may be due to its depiction of slope, rather than elevation change. In this study – as in the majority of route planning tools (Section 3.4) – colour was used to represent path slope, rather than elevation. Though height information may be derived from slope information, this process likely increased the user's cognitive load – the amount of work required to acquire, process and understand map information (Bunch & Lloyd 2006) – reducing efficiency and increasing the response time in this study.

The response time for the 2D profile was also high, albeit not significantly higher than any other visualisation. This reinforces the conclusions of Brügger (2015), who also found that 2D elevation profiles were significantly less efficient than arrow and colour visualisations, for height detection tasks. The high response time may be caused by a split in visual attention, whereby to detect the height of a point, users must match a point on the elevation profile to a point on the route (Brügger et al., 2016). Indeed, this may explain why the line width and 3D profile visualisations were most efficient for this task: both visualisations were intrinsic representations of elevation. As such, height information is located 'en-route', and next to the height markers, avoiding a so-called 'attention split' (Harrower 2007) and increasing efficiency. The low response time for the 3D profile is especially notable, given that some research has suggested the interactivity demanded by 3D visualisations to be a cause of inefficiency (Herman & Stachon, 2016).

The pattern of efficiency results was slightly different, when users performed slope detection tasks. In this instance, both 3D visualisations performed poorly: the 3D profile and 3D terrain visualisations were significantly less efficient than the 2D arrow, colour and line width visualisations. The 2D profile was also significantly less efficient than these three visualisation forms. It is presumed in the latter instance that the difference is due to the same split of visual attention as for height tasks, when matching profile to point. The high efficiency of the colour visualisation may be because, unlike with height questions, little extra cognitive processing was required to determine an answer. Instead, users were able to quickly match colour to slope, in a finding which reinforces the conclusions of Su et al. (2010). This result also strengthens a finding of Brügger (2015), who found that colour was slightly more efficient than 2D profiles and arrow visualisations when answering slope detection questions.

Similarly, the high efficiency of arrow depictions may be because users were able to quickly identify the slope steepness from clearly distinguishable arrows (1, 2, or 3 grouped arrows, each representing different slopes). By contrast, all the other visualisations (2D profile, 3D profile, 3D terrain, line-width) required an estimation of slope from line angle. It is probable that this estimation increased cognitive load, and by association response time, reducing efficiency. That line-width was significantly more efficient than the 3D profile or 3D terrain model is likely because it did not require any interaction in order for the line angle at a certain point to be viewable, or to be compared to the line angle at a different point.

There were, again, significant differences between the efficiency of different visualisations, both 2D and 3D, when answering questions related to the overall climb within a route. In this instance, the 2D route profile was one of the most efficient visualisation types – significantly more so than the 3D terrain model and colour visualisations. This is likely because this question is concerned with the entire route; as such, users were not required to match a point on the route to a point on the route profile. Thus, there was no need for a visual search, and attention was not split, increasing efficiency. Also notable was that the line-width was again one of the most efficient visualisations, and was significantly more efficient than either the 3D terrain or colour visualisations. This is due to the intrinsic nature of the visualisation, in tandem with its efficacy at conveying height information. The least efficient visualisations were, naturally, the 3D terrain and colour depictions – that one is 2D and the other 3D therefore highlights that any broad statements of 2D or 3D cartographic superiority should be treated with scepticism.

7.1.2 *Research Question 2*

Q. *Is there a significant difference in the relative effectiveness of 2D and 3D elevation visualisations for cycle route-planners, when performing route-planning tasks related to: (a) height detection; (b) slope detection; (c) overall route climb?*

In addition to highlighting efficiency differences between 2D and 3D visualisations, the study results also highlighted differences in effectiveness between these visualisation types. Again, the effectiveness of the various 2D and 3D visualisation types was task specific. When finding the highest point on a route, the arrow visualisation was least effective, by some margin. This directly contradicts the results of Brügger (2015), which found that arrow depictions were most effective at this task. Whereas in this study, the proportion of correct height answers from arrow visualisations was 19.4% (S.E. = 3.81%), in the study of Brügger (2015) the corresponding value was 87.2% (S.E. = 2.3%). The difference here may stem from the specific nature of arrow depiction used. Brügger (2015) used arrows to depict distinct 10m changes in elevation (Figure 3-6). By contrast, in this study arrows were used to depict varying degrees of path steepness, and were included for uphill slopes only (Figure 5-4). This decision was taken as the latter style is more typical of production cycling maps and route planners (c.f. Section 3.4). However, the lack of down arrows, in combination with marking steep slopes rather than segmented elevation changes, is the probable cause for the comparably low effectiveness of arrow depictions in this instance. This highlights a key issue associated with map-design research in general: overall statements regarding the efficacy of a single design type should be treated with scepticism, due to the potential variability within designs that fall under a single title, e.g. 'arrow'.

The 2D profile, whilst being one of the least efficient visualisations when answering height questions, was found to result in the highest proportion of correct answers for this task, and thus may be considered one of the most effective. The proportion of correct answers when using the 2D profile (94.4%, S.E. = 2.20%) was significantly higher than all other visualisation forms. Again, this result contradicts that of Brügger (2015), who found that 2D elevation profiles were the least effective. The 3D profile was also highly effective at this task – significantly more so than 2D arrow and colour visualisations.

Measures of the effectiveness of the visualisation types for slope tasks showed a broadly inverse pattern (Figure 6-7 vs Figure 6-8). For slope tasks, the arrow visualisation was significantly more effective than the 2D profile, 3D profile and 3D terrain visualisations, while the colour profile was also significantly more effective than the 2D and 3D profiles. By contrast, the 2D profile and 3D profile were ineffective for this task; less than 50% of questions were answered correctly using these two visualisations, compared to (S.E. = 2.66%) with the arrow visualisation. These results seem to suggest that map readers are poor judges of line slope, whether depictions are 2D or 3D. All those visualisations which used some form of line angle to convey slope (i.e. the 2D profile, 3D profile and line width, where slope) were the least effective. The results of Brügger (2015) also seem to confirm this; of the arrow, colour and 2D profile visualisations assessed in the 2015 study, the elevation profile was also found to be least effective. Further research to confirm how map readers perceive line angles would therefore be desirable.

Significant differences between the effectiveness of 2D and 3D visualisations were also revealed for questions concerning the total climb along a route. Again, it was not a simple case of either 2D or 3D visualisations reigning superior. No significant difference in effectiveness was identified between the 3D terrain, arrow and colour visualisations. Of the two most effective visualisations for this task, one was 2D (the colour visualisation) and the other 3D (the 3D profile). The colour visualisation was significantly more effective for this question type than all other visualisations, aside from the 3D profile. Tellingly, of the responses answered with this visualisation, only one was incorrect. It is feasible that participants were able to effectively (and quickly, as shown in 7.1.1) judge the overall climb from the predominant colour in the route. A mostly red or orange route immediately signals lots of climbing, while a green and yellow route indicates a largely flat path.

One curious result is that the proportion of correct answers for the 3D profile (80.6%, S.E. = 3.81%) is almost double that for the 2D profile (44.4%, S.E. = 4.78%). This is in-spite of the fact that the two visualisations show effectively the same information – one is simply in a 3D form. That the 2D profile is extrinsic should not, in this instance, make a difference: no matching between profile and route is necessary, as the user is considering the whole route. Indeed the question could theoretically be answered without a map at all. Further studies aimed at identifying the cause of the 2D vs. 3D elevation profile disparity would therefore be worthwhile.

7.1.3 *Research Question 3*

Q. *Do users prefer 2D or 3D elevation depictions when performing cycle route-planning tasks?*

User preference for visualisation type was found to vary on the basis of the particular task at hand. However, in every instance the most popular visualisation type was 2D rather than 3D. For both climb and height questions, although especially the former, the 2D elevation profile was most popular. Further insight into this preference is provided by the comments submitted by participants, a selection of which are shown in Table 7-1.

Question Type	Selected User Feedback (2D Profile)
Climb	*"In the 2D profile the user can directly see the absolute height values of the route and thus determine easily, which route has the highest amount of climbing."* *"no zooming needed"* *"Although it is tough to say where the climbs will be on the map, it is still the easiest way to see if there [are climbs] and how high the climbs are."*
Height	*"I can see clearly and faster the highest point and compare with other possible alternatives without having to explore the visualization.* *"Easy to extract the highest point, because you can just read it from the graph, instead of comparing the lines on the map. It is the only visualization that actually shows numbers."*
Slope	*"Comparison is easier and handier, because you don't have to zoom in or tilt the map."* *"I liked the 2D Profile to determine the steepest slope. No fussing required with pan tools and such. It is blatant and obvious which is the steepest route."*

Table 7-1 *Typical user feedback relating to the 2D profile, for various different question types. A complete collection of all user feedback is included in Appendix 3.*

Thus, users appeared to prefer the 2D profile because it required less interaction, allowed easier comparison and they felt as though it gave more precise answers – even if the effectiveness results suggest otherwise. Where the 2D profile falls down in terms of preference, it is for the lack of dynamic brushing that would allow points on the map to be directly related to points on the profile. Indeed, several respondents cite this as the reason they did not choose the 2D profile as their preferred visualisation type. Brügger (2015) also found that elevation profiles were most preferred by users, although only for questions of slope and route selection. For questions of height, the arrow depiction was preferred (Brügger, 2015). By contrast, in this study, no participants selected the arrow depiction as their favourite visualisation for answering height questions. This difference is likely due to the disparity in style of arrow depictions, as discussed in Section 7.1.2.

In general, users appeared to dislike the 3D visualisations. For climb tasks the 3D profile was least popular (preferred by only one study participant), while for slope tasks the 3D terrain was the least popular (preferred by 8.3% of study participants). From the comments, this dislike of the 3D visualisations primarily stemmed from the need for interactivity, and the necessity to 'explore' the visualisations in order to answer the questions. One participant, who preferred the 2D profile, simply stated "I don't like handling the 3D visualizations." By contrast, other research, has shown a general favour for 3D over 2D visualisation (e.g. Popelka and Brychtova, 2013; Wilkening and Fabrikant, 2011), or for more 'realistic', pseudo-3D maps (e.g. Schobsberger and Patterson, 2008; Wilkening and Fabrikant, 2011). However, these studies have all used static rather than interactive 3D depictions. The need to dynamically manipulate the maps in order to answer the questions may account for the unpopularity of 3D depictions in this study. Indeed, participant comments such as, "If I have plenty of time I would prefer a 3D visualisation" would seem to reinforce this inference.

That is not to say, however, that the 3D visualisations were universally despised. For height questions, for example, 25% of users preferred the 3D profile visualisation. In this instance, users commented that the location of the profile along the path, avoided the difficulty of matching path with profile (the same issue that would be negated by interactive brushing). Further, as interactive and 3D maps become more ubiquitous, and users become more used to navigating 3D environments, it may be that more users begin to prefer 3D visualisations.

The results of the study broadly supported the conclusion of much other research, which has found that users often do not prefer those cartographic designs which are most efficient and effective (Hegarty, 2009; Brügger, 2015). This is particularly true for the 2D elevation profile, albeit not for height tasks, as is discussed above. However, the same trend is apparent with other visualisation forms. For example, the line-width visualisation was, on average, both efficient and effective – yet received consistently low preference ratings from study participants. This may be because the line-width visualisation was unfamiliar, having only been used within the academic paper of Huffman (2009). However, this can not be stated for certain, as questions about familiarity with the different visualisation types were not asked before the study, and while some work suggests that familiarity breeds preference (Hegarty, 2009), other research suggests the opposite (Werner 1993).

7.2 *Additional Findings and Wider Implications*

Though the primary focus of this study lay on the relative usability of 2D and 3D visualisations, it is also possible to draw a number of wider conclusions from the results. One key finding, which supports previous work by others (Liao et al. 2016; Schobesberger & Patterson 2008; Brügger et al. 2016; Savage et al. 2004) is confirmation that the usability of different visualisations is heavily dependent on the task at hand. This statement appears to hold true for all three components of usability – efficiency, effectiveness and user preference. Thus, when designing future cycleling maps and cycle route-planners, a crucial question must be to ask *exactly* what users need, and what questions are they using the cycling map or cycle route-planner to answer. Only when this has been determined, can the most usable and desirable visualisation type be chosen.

Throughout the literature, the terms '3D cartography', '3D mapping', '3D maps', and so forth, appear to be used largely interchangeably. In many instances, the same term may be used to refer to vastly different forms of cartography. Crucially, the preceding results show that the usability of 3D cartographic visualisations is highly variable, depending on the exact form that the '3D map' is taking. This finding highlights the need to be critical about broad-brush statements concerning the usability of '3D cartography'. Always, the question of 'what kind of 3D cartography is being referred to?' should be asked.

Beyond assessing the relative usability of a range of 2D and 3D terrain depictions for cycling maps, this study was also the first to empirically test the line-width visualisation proposed by Huffman (2009). Although the users tested within this study generally disfavoured this visualisation type, the line-width visualisation did show promise. Notably, this visualisation form was consistently one of the most efficient, and often one of the most effective, for all different task types. With this depiction, a user can at a glance see absolute elevation, slope, and total climb for the overall route. Their attention is not split, making the visualisation more efficient than a 2D profile, while the lack of interaction means that response times are generally faster than for 3D visualisations. Thus, despite its sometimes low popularity, line-width appears a promising means of communicating the terrain along a path, and should be explored further in future cycle route-planners.

8 Study Limitations and Avenues of Future Research

Although the study provided firm results regarding the relative usability of a range of 2D and 3D terrain visualisations for cycle route-planners, the findings must be viewed in the context of certain methodological limitations, which are discussed below. Potential avenues of future research – aimed both at addressing these weaknesses, and also at advancing the field of map-design research in general – are presented alongside.

Though every effort was made to form a rigorous methodology, some key issues became apparent only during the analysis stage. Crucially, while the methodology presented in Chapter 5 was aimed at alleviating learning effects, its design was such that subsequent data analysis had to use data from all participants simultaneously. Consequently, it was not possible to determine the influence of factors such as educational background or gender on the overall results. If analysis was conducted on the basis of gender, for example, it would not be possible to ensure that both groups (i.e. male and female) were exposed to the same combinations of visualisation and spatial area scenarios. Thus, any differences in usability may stem from these differences, rather than gender or the inherent usability of the visualisations. An improved methodological design which counters this problem, but also avoids the learning effects observed in Brügger et al (2016) would therefore be desirable for future studies.

Additional criticism may be levelled at the particular forms of terrain visualisation which were tested in this study. In the study of Brügger (2015), visualisations were designed such that they all displayed changes in relative or absolute height (c.f. Section 5.2.1.2). However, in some instances this led to the use of visualisations, in particular the arrow design (Figure 3-6) which are wholly atypical of 'real-life' visualisations. The wider applicability of any related conclusions is thusly cast in doubt. By contrast, the designs in this study were chosen to as closely resemble published visualisations as possible. Problematically, this means that while some designs show only slope (e.g. the colour visualisation), others show slope and height (e.g. the 2D profile). As such, it is arguable that some differences in usability may stem not from intrinsic design characteristics, but instead differences in the data that the designs were representing. Countering this problem, whilst simultaneously staying 'true' to real-life designs, is challenging, but would be a worthy goal of future research. For example, within the broad 2D vs. 3D debate, more reliable conclusions regarding relative usability may be derived, if it was ensured that tested designs were more directly comparable. For example, a 2D elevation profile should be compared to a 3D elevation profile, a continuous 2D contour map should be compared to a continuous 3D terrain map, and so on. While the downside of this approach is that the number of available comparisons would be reduced, subsequent conclusions would arguably be of greater validity.

An additional limitation of the study was the omission of eye movement recording. This choice was taken due to the difficulties of using the technique with interactive 3D visualisations (Section 3.2). However, Brügger (2015) found that eye tracking data was crucial in understanding usability differences between visualisation types. Although in the early stages, some recent research has also begun to develop methods for using eye-tracking data with interactive applications (Ooms et al., 2014). If such techniques could be applied to a repeat of this study, further insight into the causes behind usability differences could be uncovered.

A notable issue with the user study was also the demographic and educational background of the participants. The majority of participants were aged 21-30, were highly educated, and had strong pre-existing cartographic knowledge (Section 6.1). As noted, the study methodology meant it was not possible to analyse the data on the basis of educational background. Whether or not the strong trend of pre-existing cartographic knowledge influenced the results can therefore not be concluded with certainty. However, it is quite likely that the experience of the study participants did indeed influence the measured usability criteria, or introduce a certain bias into the results. For example, the strong technical background of the majority of users (i.e. studying at a Technical University) may have improved the performance of the interactive visualisations, as these users are likely more comfortable with such tools than the average user. Likewise, these users' experience of map-reading likely further improved the efficiency and effectiveness of the 'traditional' 2D elevation representations, with which they are likely familiar. However, these statements are merely conjecture, and the exact influence of the participants' background remains uncertain. A repeat of the study with a wider variety of participants – including a higher proportion with no cartographic training – would help clarify this matter.

Both this study, and the comparable work of Brügger (2015), were conducted in highly controlled, 'idealised' environments. Light and noise levels were constant, and the participants remained uninterrupted and unhindered whilst completing the assessment. Though this setup helps isolate the variables we wish to study – that is, the usability of various 2D and 3D visualisations – it reduces the degree of realism, and thus the wider applicability of the results. There is therefore an argument for conducting a similar study in a more 'realistic' environment, to see if the results were comparable. Indeed, given the increasing ubiquity of mobile cartography, and the growth in options for field-based mobile assessment (Burghardt & Wirth 2011), it would be especially desirable to conduct a similar study in the field, using mobile devices and visualisations optimised for mobile displays.

9 Conclusion

Building on the work begun by Brügger et al. (2016), this thesis has begun to expand the fairly limited body of research concerning the design of cycling maps. In particular, by adopting a methodology that allowed empirical testing of interactive visualisations within a controlled environmental setting, it has permitted comparison of the usability of a range of 2D and 3D elevation depictions for cycle route-planners. The variability between different forms of 2D and 3D terrain visualisations means that it is impossible to state conclusively, or even broadly, that the added dimension either improves or reduces usability. However, it remains possible to draw some smaller-scale conclusions from the results of the user-study. With especial reference to the research questions which drove the study, the following conclusions may be stated:

1. The relative efficiency of 2D and 3D terrain visualisations depends on the particular cycle route-planning task for which they are being used. However, the 2D line-width visualisation was notably efficient in the widest range of use-scenarios. Depending on the task at hand, the 2D arrow and colour visualisations were also very efficient. By contrast, the 3D terrain model was consistently inefficient, and always less efficient than the 3D elevation profile. Whilst these results may partially be a reflection of pre-existing user-familiarity with the different visualisation forms, the methodology utilised in the study did not allow testing of this hypothesis.

2. Again, the relative efficacy of 2D and 3D terrain visualisations was found to depend upon the particular cycle route-planning task for which they are being used. As such, it is not feasible to widely state that one dimensionality is more effective than another. However, broadly speaking, the results appeared to show that the lower the cognitive load a visualisation placed on the user, the more effective it was at communicating terrain information –

this confirms earlier research related to the cognitive load of cartography (Bunch & Lloyd 2006). Indeed, this factor appears to have had a greater influence on efficacy than the dimensionality of the visualisations.

3. The majority of users preferred 2D visualisations, for all cycle route-planning tasks tested. In particular, the 2D elevation profile was especially popular amongst test participants. 3D visualisations were generally disliked, primarily due to the necessity for, and perceived complexity of, user interaction. Notably, the 3D elevation profile appeared consistently more popular than 3D terrain models. Whether or not user preferences toward 3D depictions would change with alternative interaction forms, or indeed increased user familiarity, remains to be seen.

Cumulatively, these results paint a varied picture for the future of interactive 3D visualisations in cycle route-planning, and indeed in other route-planning and cartographic situations. However, the results appear to show that more abstracted, innovate 3D cartographic depictions present a more usable option than the more frequently studied 3D terrain models. Perhaps this result was to be expected; if maps are useful because they abstract a complex reality, then cartographies which attempt to reproduce that reality arguably do so in contradiction to one of the great strengths of traditional cartography.

10 References

Adams, P.C., 2009. Geographies of Media and Communication: A Critical Introduction., Oxford: Wiley-Blackwell.

Akerman, J.R., 2002. American Promotional Road Mapping in the Twentieth Century. Cartography and Geographic Information Science, 29(3), pp.175–192. Available at: http://www.tandfonline.com/doi/abs/10.1559/152304002782008459.

Bevan, N., Carter, J. & Harker, S., 2015. ISO 9241-11 revised: What have we learnt about usability since 1998? In M. Kuruso, ed. Human Computer Interaction. pp. 143–151. Available at: http://www.nigelbevan.com/papers/ISO 9241-11 revised- What have we learnt about usability since 1998.pdf.

Bike Hike, 2017. Bike Hike - Course Creator. Available at: http://www.bikehike.co.uk [Accessed November 14, 2016]

Bleisch, S. & Dykes, J., 2006. Planning Hikes Virtually–How Useful are Web- based 3D Visualizations? In Proceedings GIS Research UK 14th Annual Conference. Nottingham, UK, pp. 313–318.

Bleisch, S. & Dykes, J., 2008. Using Web-Based 3-D Visualization for Planning Hikes Virtually - An Evaluation. In N. J. Mount et al., eds. Representing, Modeling and Visualizing the Natural Environment: Innovations in GIS 13. Florida: CRC Press, pp. 353–365.

Boer, A. et al., 2012. Abstracting the Reality: Usability Evaluation of Levels of Abstraction and Realism in Geographic Visualizations. University of Zurich.

Brodersen, L., 2001. Maps as Communication: Theory and Modelling in Cartography First Edit., National Survey and Cadastre Denmark.

Brügger, A., 2015. Where are the Ups and Downs? Evaluating Elevation Representations for Bicycle Paths in City Maps. University of Zürich.

Brügger, A., Fabrikant, S.I. & Çöltekin, A., 2016. An empirical evaluation of three elevation change symbolization methods along routes in bicycle maps. Cartography and Geographic Information Science, 406(September), pp.1–16.

Buchroithner, M.F. & Knust, C., 2013. True-3D in cartography – current hard and softcopy developments. In M. A & I. Drecki, eds. Geospatial Visualisation. Heidelberg: Springer, pp. 41–65.

Bunch, R.L. & Lloyd, R., 2006. The Cognitive Load of Geographic Information. The Professional Geographer, 58(October 2004), pp.209–220. Available at: http://www.tandfonline.com/doi/abs/10.1111/j.1467-9272.2006.00527.x.

Burghardt, D. & Wirth, K., 2011. Comparison of Evaluation Methods of Field-Based Usability Studies. In Proc. International Cartographic Conference (2011). p. 6.

Castner, H.W. & Wheate, R., 1979. Re-assessing the Role Played by Shaded Relief in Topographic Scale Maps. The Cartographic Journal, 16(2), pp.77–85.

Ciolkosz-Styk, A., 2012. The visual search method in map perception research. Geoinformation Issues, 4(1), pp.33–42. Available at: http://bc.igik.edu.pl/Content/357/GI_2013_5.pdf.

City of Vancouver, 2016. Vancover Cycling Map. Cycling Map – City of Vancouver. Available at: http://vancouver.ca/files/cov/map-cycling-vancouver.pdf [Accessed December 12, 2016].

Cohen, C.A. & Hegarty, M., 2007. Individual Differences in Use of External Visualizations to Perform an Internal Visualization Task. Applied Cognitive Psychology, 21, pp.701–711. Available at: http://

www.scopus.com/inward/record.url?eid=2-s2.0-37349131250&partnerID=40&md5=b2db461a1b1b0e6fe1ad3312b05a0329.

Collier, P., Forrest, D. & Pearson, A., 2003. The Representation of Topographic Information on Maps: The Depiction of Relief. The Cartographic Journal, 40(1), pp.17–26. Available at: http://www.tandfonline.com/doi/full/10.1179/000870403235002033.

Çöltekin, A. et al., 2009. Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis. Cartography and Geographic Information Science, 36(1), pp.5–17.

Crampton, J.W., 2001. Maps as social constructions: power, communication and visualization. Progress in Human Geography, 25(2), pp.235–252.

Crampton, J.W. & Krygier, J., 2005. An introduction to critical cartography. ACME, 4(1), pp.11–33.

Crawford, P. V., 1973. The Perception of Graduated Squares as Cartographic Symbols. The Cartographic Journal, 10(2), pp.85–88.

Dickinson, E.A., 2012. How Bicycle Maps and Trip Planners Can Represent Experience. Simon Fraser University.

Edney, M.H., 2005. Putting "Cartography" into the history of cartography: Arthur H. Robinson, David Woodward, and the creation of a discipline. Cartographic Perspectives, (51), pp.14–29.

Ehlers, M., Jung, S. & Stroemer, K., 2002. Design and Implementation of a GIS Based Bicycle Routing System for the World Wide Web (WWW). In Symposium on Geospatial Theory, Processing and Applications. Ottawa: International Society for Photogrammetry and Remote Sensing, p. 5. Available at: http://www.isprs.org/proceedings/XXXIV/part4/pdfpapers/198.pdf.

Ekman, G. & Junge, K., 1961. Psychophysical Relations in Visual Perception of Length, Area and Volume. Scandinavian Journal of Psychology, 2(1), pp.1–10.

Ekman, G., Lindman, R. & William-Olsson, W., 1963. A Psychophysical Study of Cartographic Symbols. Geografiska Annaler, 45(4), pp.262–271.

Environment Agency, 2016. LIDAR Composite DSM- 1m. data.gov.uk. Available at: https://data.gov.uk/dataset/lidar-composite-dtm-1m1 [Accessed December 18, 2016].

Farkas, D.K., 2013. Evaluation and Usability Testing. University of Washington – Dept. of Human Centered Design & Engineering, p.7. Available at: http://faculty.washington.edu/farkas/HCDE 407-2013/Evaluation and Usability Testing.pdf [Accessed November 14, 2016].

Field, A.P., 2009. Discovering statistics using SPSS (and sex, drugs and rock'n'roll), 3rd Edition, Sage Publications Ltd., London.

Flannery, J.J., 1956. The graduated circle: A description, analysis, and evaluation of a quantitative map symbol. University of Wisconsin, Madison.

Frøkjær, E., Hertzum, M. & Hornbæk, K., 2000. Measuring Usability : Are Effectiveness , Efficiency , and Satisfaction Really Correlated ? ACM CHI 2000 Conference on Human Factors in Computing Systems, 2(1), pp.345–352.

Griffin, T.L.C., 1985. Group and Individual Variations in Judgment and Their Relevance To the Scaling of Graduated Circles. Cartographica: The International Journal for Geographic Information and Geovisualization, 22(1), pp.21–37. Available at: http://utpjournals.press/doi/10.3138/0X46-G750-6261-R826.

Griffin, T.L.C. & Lock, B.F., 1979. (1979) The Perceptual Problem in Contour Interpretation.pdf. The Cartographic Journal, 16(2), pp.61–71.

Harrower, M., 2007. The Cognitive Limits of Animated Maps. Cartographica: The International Journal for Geographic Information and Geovisualization, 42(4), pp.349–357.

Hegarty, M. et al., 2009. Naïve Cartography: How Intuitions about Display Configuration Can Hurt Performance. Cartographica: The International Journal for Geographic Information and Geovisualization, 44(3), pp.171–186. Available at: http://utpjournals.metapress.com/openurl.asp?genre=article&id=doi:10.3138/carto.44.3.171.

Heim, F., 2014. A Visual Search Efficiency Study: An Evaluation of Labels, Road Junctions and Landmarks in 2D Orthogonal Maps.

Herlihy, D. V, 2004. Bicycle: The History. New Haven: Yale University Press.

Herman, L. & Stachon, Z., 2016. Comparison of User Performance With Interactive and Static 3D Visualization – Pilot Study. In ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Prague: 2016 XXIII ISPRS Congress, 12–19 July 2016, pp. 655–661. Available at: http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B2/655/2016/isprs-archives-XLI-B2-655-2016.pdf.

Hochmair, H., 2004. Decision Support for Bicycle route-planning In Urban Environments. In 7th AGILE Conference on Geographic Information Science" 29 April-1May 2004, Heraklion, Greece Parallel Session 8.2- "Decision Support Systems. Heraklion, Greece: Crete University Press, pp. 697–706.

Hochmair, H., 2005. Towards a classification of route selection criteria for route planning tools. In P. Fisher, ed. Developments in Spatial Data Handling. Berlin: Springer, pp. 481–92.

Holzinger, A., 2005. Usability {Engineering} {Methods} for {Software} {Developers}. Communications of the ACM, 48(1), pp.71–74. Available at: http://doi.acm.org/10.1145/1039539.1039541.

Huffman, D., 2009. A Technique for Encoding Elevation Changes Along a Route. Cartographic Perspectives, 63, pp.83–86.

Incoul, A., Ooms, K. & De Maeyer, P., 2015. Comparing Paper and Digital Topographic Maps Using Eye Tracking. In J. Brus, A. Vondraková, & V. Vozenílek, eds. Modern Trends in Cartography. Springer, pp. 339–356. Available at: http://link.springer.com/10.1007/978-3-319-07926-4.

Irvankoski, K.M., 2012. Visualisation of Elevation Information on Maps : an Eye Movement Study. University of Helsinki.

Ivory, M.Y., 2001. Usability evaluation of user interfaces. In An Empirical Foundation for Automated Web Interface Evaluation. Berkeley, California: UC Berkeley Computer Science Division, pp. 4–53.

Kalt, B., 2015. The influence of visualisation type on decision-making: A comparative evaluation of 2D, 3D and combined 2D/3D geovisualisations for urban planning decisions. University of Zurich.

Kitchin, R., Perkins, C. & Dodge, M., 2009. Thinking About Maps. In M. Dodge, C. Perkins, & R. Kitchin, eds. Rethinking Maps. Abingdon: Routledge, pp. 1–32.

Kolácný, A., 1969. Cartographic Information-a Fundamental Concept and Term in Modern Cartography. The Cartographic Journal, 6(1), pp.47–49.

Krygier, J., 2007. Perceptual Scaling of Map Symbols. Making Maps : DIY Cartography – Resources and Ideas for Making Maps. Available at: https://makingmaps.net/2007/08/28/perceptual-scaling-of-map-symbols/ [Accessed November 3, 2016].

Kulhavy, Raymond W; Pridemore, Doris R; Stock, W.A., 1992. Cartographic Experience And Thinking Aloud About Thematic Maps. Cartographica, 29(1), pp.1–9. Available at: http://www.utpjournals.press/doi/abs/10.3138/H61J-VX35-J6WW-8111?journalCode=cart.

Liao, H. et al., 2016. Exploring differences of visual attention in pedestrian navigation when using 2D maps and 3D geo-browsers. Cartography and Geographic Information Science, pp.1–17.

Liu, P., 2015. Plan a ride with Surface, Directions, and Turf.js. Mapbox. Available at: https://www.mapbox.com/blog/dc-bikeshare-revisited/ [Accessed October 10, 2016].

MacEachren, A.M., 2013. Cartography as an Academic Field: A Lost Opportunity or a New Beginning? Cartographic Journal, 50(2), pp.166–170.

MacEachren, A.M., 1994. Visualization in modern cartography: Setting the agenda. In A. M. MacEachren, ed. Visualization in Modern Cartography. London: Pergamon, pp. 1–13.

MacEachren, A.M. & Kraak, M.J., 1997. Exploratory Cartographic Visualisation: Advancing the Agenda. Computer & Geosciences Vol. 23, 23(4), pp.335–343.

Maibach, E., Steg, L. & Anable, J., 2009. Promoting physical activity and reducing climate change: Opportunities to replace short car trips with active transportation. Preventive Medicine, 49(4), pp.326–327. Available at: http://dx.doi.org/10.1016/j.ypmed.2009.06.028.

Manson, S.M. et al., 2012. Using Eye-tracking and Mouse Metrics to Test Usability of Web Mapping Navigation. Cartography and Geographic Information Science, 39(1), pp.48–60.

McDonald, J.H., 2014. Handbook of Biological Statistics. Sparky House Publishing, pp.59–67.

Montello, D.R., 2002. Cognitive Map-Design Research in the Twentieth Century: Theoretical and Empirical Approaches. Cartography and Geographic Information Science, 29(3), pp.283–304.

Movingworld GMBH, 2016. Maps 3D Pro. Apple iTunes Store. Available at: https://itunes.apple.com/at/app/maps-3d-pro-gps-fur-fahrrad/id391304000?mt=8 [Accessed December 19, 2016].

Nicholson, T., 2004. Cycling and Motoring Maps in Western Europe 1885–1960. The Cartographic Journal, 41(3), pp.181–215. Available at: http://www.tandfonline.com/doi/full/10.1179/000870404X17748.

Niedomysl, T. et al., 2013. Learning Benefits of Using 2D Versus 3D Maps: Evidence from a Randomized Controlled Experiment. Journal of Geography, 112(3), pp.87–96. Available at: http://www.tandfonline.com/doi/abs/10.1080/00221341.2012.709876.

Nielsen, J., 1993. Usability Engineering, Available at: http://www.useit.com/jakob/useengbook.html.

Nivala, A., 2005. User-Centred Design in the Development of a Mobile Map Application. Univeristy of Helsinki. Available at: http://www.soberit.hut.fi/T-121/shared/thesis/Nivala_LicentiateThesis.pdf.

Norman, D.A., 2002. The Design of Everyday Things, New York: Basic Books. Available at: http://ucdwiki.chuank.com/uploads/Main/UCDReading_wk5.pdf.

Olson, J.M., 2005. Arthur H. Robinson and the Fabric of Cartography. In International Cartographic Conference. A Coruña, Spain: ICAI, pp. 1–7.

Ooms, K. et al., 2014. Combining user logging with eye tracking for interactive and dynamic applications. Behavior Research Methods, 45, pp.977–993. Available at: http://link.springer.com/10.3758/s13428-014-0542-3.

Ordnance Survey, 2016. OS Open Data Supply. Mapping data and geographic information from OS. Available at: https://www.ordnancesurvey.co.uk/opendatadownload/products.html [Accessed December 15, 2016].

Patterson, T. & Jenny, B., 2011. The Development and rationale of cross-blended hypsometric tints. Cartographic Perspectives, (69), pp.31–45.

Petchenik, B.B., 1983. A mapmaker's perspective on map design research 1950-1980. In D. R. F. Taylor, ed. Graphic communication and design in contemporary cartography. Chicester, UK: John Wiley & Sons, pp. 37–68.

Phillips, R.J., 1982. An experimental investigation of layer tints for relief maps in school atlases. Ergonomics, 25(12), pp.1143–1154.

Phillips, R.J., 1979. An Experiment with Contour Lines.pdf. The Cartographic Journal, 16(2), pp.72–76.

Phillips, R.J., De Lucia, A. & Skelton, N., 1975. Some objective tests of the legibility of relief maps. The Cartographic Journal, 12(1), pp.39–46.

Phillips, R.J. & Noyes, L., 1978. An objective comparison of relief maps produced with the SYMAP and SYMVU programs. Bulletin of the Society of University Cartographers, 12, pp.13–25.

Popelka, S. & Brychtova, A., 2013. Eye-tracking Study on Different Perception of 2D and 3D Terrain Visualisation. The Cartographic Journal, 50(3), pp.240–246. Available at: http://openurl.ingenta.com/content/xref?genre=article&issn=0008-7041&volume=50&issue=3&spage=240.

Popelka, S. & Dolezalova, J., 2015. Non-photorealistic 3D Visualization in City Maps: An Eye-Tracking Study. In J. Brus, A. Vondraková, & V. Vozenílek, eds. Modern Trends in Cartography. Springer, pp. 357–367. Available at: http://link.springer.com/10.1007/978-3-319-07926-4.

Rautenbach, V. et al., 2015. An Assessment of Visual Variables for the Cartographic Design of 3D Informal Settlement Models. In Proceedings of the ICC 2015: Rio de Janeiro, Brazil. p. 15.

Relive, 2016. Relive.cc. Available at: https://www.relive.cc [Accessed December 17, 2016].

Robinson, A.H., 1952. The Look of Maps: An Examination of Cartographic Design, Madison: University of Wisconsin Press.

Robinson, A.H. & Petchenik, B.B., 1975. The map as a communication system. The Cartographic Journal, 12(1), pp.7–15. Available at: http://utpjournals.metapress.com/index/023561171TW606T5.pdf.

Roth, R.E., 2013. Cartographic Interaction: What we know and what we need to know. Journal of Spatial Information Science, 6(6), pp.59–115.

Rubinstein, D., 1977. Cycling in the 1890s. Victorian Studies, 21(1), pp.47–71.

Savage, D.M., Wiebe, E.N. & Devine, H. a., 2004. Performance of 2D versus 3D Topographic Representations for Different Task Types. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 1793–1797. Available at: http://pro.sagepub.com/lookup/doi/10.1177/154193120404801601.

Scharfe, W., 1986. Max Eckert's " Kartenwissenschaft ". The Turning Point in German Cartography. Imago Mundi, 38, pp.61–66.

Schobesberger, D. & Patterson, T., 2008. Evaluating the effectiveness of 2D VS. 3D trailhead maps: A map user study conducted at Zion National Park, United States. Bulletin of the Society of Cartographers, 42(1–2), pp.3–8.

Seattle Department of Transportation, 2016. 2016 Bike Map Brochure. Available at: http://www.seattle.gov/transportation/docs/2016_BikeMap_Brochure.pdf [Accessed December 12, 2016].

Shepherd, I., 2007. Travails in the Third Dimension : a critical evaluation of 3D geographical visualization. In M. Dodge & M. Turner, eds. GeoVisualisation. London: Wiley, pp. 199–222.

Sherwin, H. and Bartle, C. (2012) Evaluating cycle mapping styles. Project Report. University of the West of England. Available from: http://eprints.uwe.ac.uk/16823

Sherwin, H. & Melia, S., 2012. Cycle Mapping in the UK and the "London Cycle Map," Bristol. Available at: http://eprints.uwe.ac.uk/17899/.

Shurtleff, M. & Geiselman, E., 1986. A Human-Performance Based Evaluation of Topographic Maps and Map Symbols with Novice Map Users. The Cartographic Journal1, 23, pp.52–55.

Slocum, T.A. et al., 2005. Thematic Cartography and Geovisualization, Upper Saddle River: Pearson Prentice Hall. Available at: http://www.amazon.com/Thematic-Cartography-Geovisualization-3rd-Edition/dp/0132298341.

Smallman, H.S. & John, M.S., 2005. Naive realism: misplaced faith in realistic displays. Ergonomics in Design: The Quarterly of Human Factors Applications, 13(3), pp.6–13. Available at: http://erg.sagepub.com/content/13/3/6.short.

Speich, D., 2009. Mountains Made in Switzerland: Facts and Concerns in Nineteenth-Century Cartography. Science in Context, 22(3), p.387. Available at: http://www.journals.cambridge.org/abstract_S0269889709990068.

St John, M. et al., 2001. The Use of 2D and 3D Displays for Shape-Understanding versus Relative-Position Tasks. Human Factors: The Journal of the Human Factors and Ergonomics Society, 1(43), pp.79–98.

Steinke, T.R., 1987. Eye Movement Studies In Cartography And Related Fields. Cartographica: The International Journal for Geographic Information and Geovisualization, 24(2), pp.40–73. Available at: http://utpjournals.press/doi/10.3138/J166-635U-7R56-X2L1.

Su, J.G. et al., 2010. Designing a route planner to facilitate and promote cycling in Metro Vancouver, Canada. Transportation Research Part A: Policy and Practice, 44(7), pp.495–505. Available at: http://dx.doi.org/10.1016/j.tra.2010.03.015.

Tchoukanski, I., 2016. ET Surface 6.2. ET Surface. Available at: http://www.ian-ko.com/ET_Surface/ets_main.htm [Accessed January 17, 2017].

Werner, R.J., 1993. A Survey of Preference Among Nine Equator-Centered Map Projections. Cartography and Geographic Information Science, 20(1), pp.31–39.

Wessel, N. & Widener, M., 2015. Rethinking the urban bike map for the 21st century. Cartographic Perspectives, 2015(81), pp.6–22.

Whitefield, A., Wilson, F. & Dowell, J., 1991. A Framework for Human Factors Evaluation. Behaviour & Information Technology, 10(1), pp.65–79.

Wilkening, J. & Fabrikant, S.I., 2011. How do decision time and realism affect map-based decision making? In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 1–19.

Wilkening, J. & Fabrikant, S.I., 2013. How users interact with a 3D geo-browser under time pressure. Cartography and Geographic Information Science, 40(1), pp.40–52.

Winters, M. et al., 2011. Motivators and deterrents of bicycling: Comparing influences on decisions to ride. Transportation, 38(1), pp.153–168.

Worth, C., 1989. Problems with experimental research in map design - a case study. The Cartographic Journal, 26(2), pp.148–153.

Zyszkowska, W., 2015. Map perception: theories and research in the second half of the twentieth century. Polish Cartographical Review, 47(4), pp.179–190. Available at: http://www.degruyter.com/view/j/pcr.2015.47.issue-4/pcr-2015-0017/pcr-2015-0017.xml.

11 Appendix

1. Questionnaire Example (Group 1)

In total, 6 different versions of the following questionnaire were produced, using the combination of visualisations and spatial areas noted in Table 5-2. Due to the length of each questionnaire, and their broad structural similarity, only one example – for Group 1 – is included for reference below.

0%

Pre-Questionnaire

Before we begin, we'd like to collect some basic background information about yourself. All this information is treated confidentially, and stored anonymously.

Please enter your personal ID (given to you by the instructor):

✱ How old are you?

❶ Choose one of the following answers

- Under 16
- 16-20
- 21-30
- 31-40
- 41-50
- 51-60
- 61-70
- 71-80
- Over 80

✱ Please select your gender.

♀ Female	♂ Male

✱ On average, how often do you use a bicycle?

❶ Choose one of the following answers

- Daily
- Several times weekly
- Once weekly
- Monthly
- Yearly
- Never

✱ How would you rate your training and expertise in the following?

	1 (no knowledge)	2	3	4	5 (expert)
Cartography and GIS	1 (no knowledge)	2	3	4	5 (expert)
Geography	1 (no knowledge)	2	3	4	5 (expert)
Graphic Design & Fine Arts	1 (no knowledge)	2	3	4	5 (expert)
3D Graphics	1 (no knowledge)	2	3	4	5 (expert)
User-Centred Design	1 (no knowledge)	2	3	4	5 (expert)

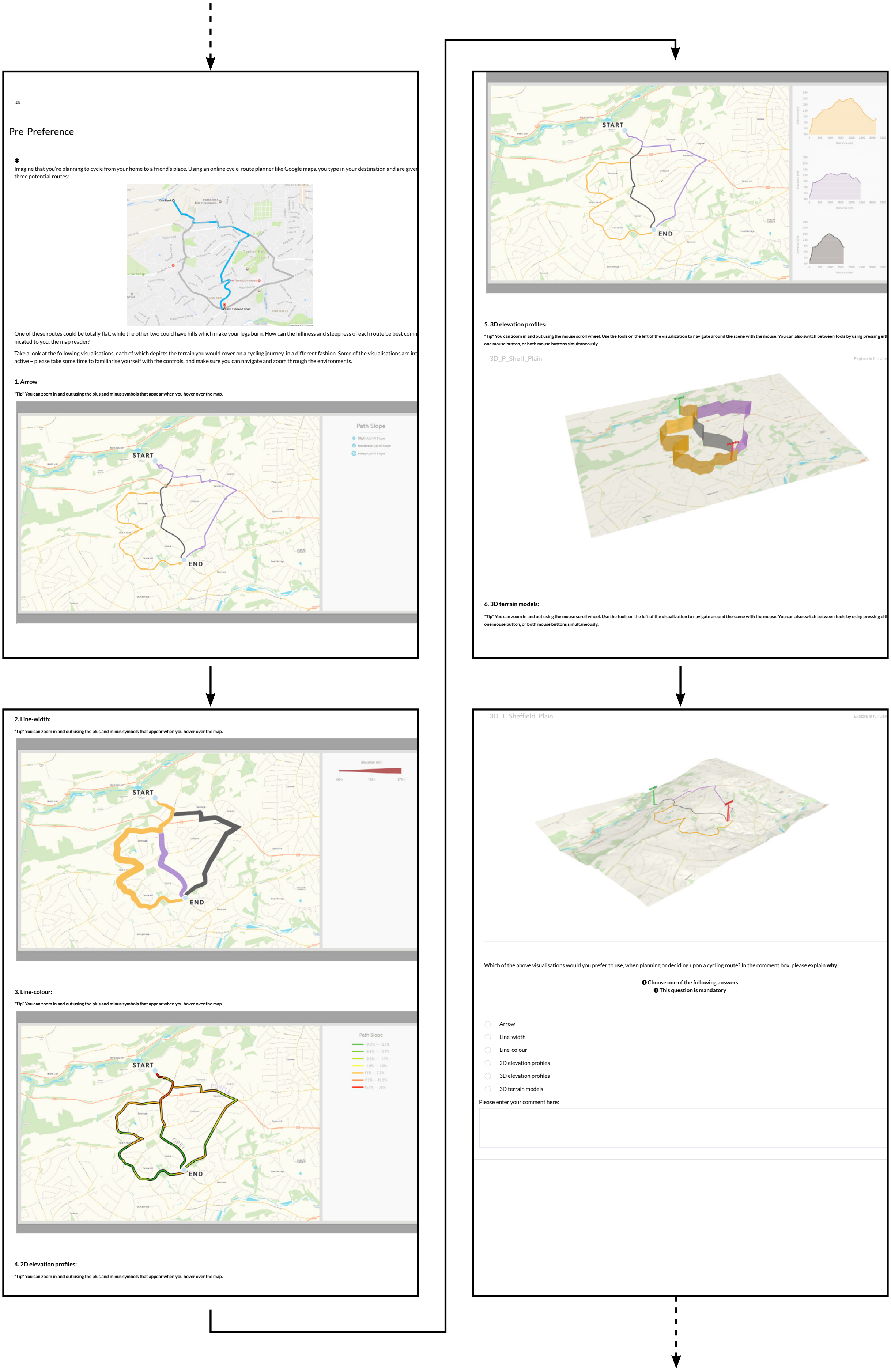
✱ How would you rate your experience and knowledge in the following?

	1 (no knowledge)	2	3	4	5 (expert knowledge)
3D Graphics	1 (no knowledge)	2	3	4	5 (expert knowledge)
Graphics of any kind (e.g. charts, graphs, photos etc.)	1 (no knowledge)	2	3	4	5 (expert knowledge)
Spatial data and geographic visualisations of any kind (e.g. maps, terrain models, satellite imagery etc.)	1 (no knowledge)	2	3	4	5 (expert knowledge)
The English language	1 (no knowledge)	2	3	4	5 (expert knowledge)

✱ How would you rate your experience and knowledge in the following?

	1 (no knowledge)	2	3	4	5 (expert knowledge)
Map reading (paper maps)	1 (no knowledge)	2	3	4	5 (expert knowledge)
Map reading (mobile maps and apps)	1 (no knowledge)	2	3	4	5 (expert knowledge)
Interactive map applications	1 (no knowledge)	2	3	4	5 (expert knowledge)
3D map applications	1 (no knowledge)	2	3	4	5 (expert knowledge)

✱ What is your highest attained level of education?


2%
Pre-Preference
*
Imagine that you're planning to cycle from your home to a friend's place. Using an online cycle-route planner like Google maps, you type in your destination and are giver
three potential routes:
Den Bank
400 Fulwood Road
One of these routes could be totally flat, while the other two could have hills which make your legs burn. How can the hilliness and steepness of each route be best comm
nicated to you, the map reader?
Take a look at the following visualisations, each of which depicts the terrain you would cover on a cycling journey, in a different fashion. Some of the visualisations are int
active – please take some time to familiarise yourself with the controls, and make sure you can navigate and zoom through the environments.
1. Arrow
Tip You can zoom in and out using the plus and minus symbols that appear when you hover over the map.
START
END
Path Slope
Slight Uphill Slope
Moderate Uphill Slope
Steep Uphill Slope
2. Line-width:
Tip You can zoom in and out using the plus and minus symbols that appear when you hover over the map.
START
END
Elevation (m)
3. Line-colour:
Tip You can zoom in and out using the plus and minus symbols that appear when you hover over the map.
START
END
Path Slope
4. 2D elevation profiles:
Tip You can zoom in and out using the plus and minus symbols that appear when you hover over the map.
START
END
Elevation (m)
Distance (m)
5. 3D elevation profiles:
Tip You can zoom in and out using the mouse scroll wheel. Use the tools on the left of the visualization to navigate around the scene with the mouse. You can also switch between tools by using pressing eit
one mouse button, or both mouse buttons simultaneously.
3D_P_Sheff_Plain
Explore in full view
START
FINISH
6. 3D terrain models:
Tip You can zoom in and out using the mouse scroll wheel. Use the tools on the left of the visualization to navigate around the scene with the mouse. You can also switch between tools by using pressing eit
one mouse button, or both mouse buttons simultaneously.
3D_T_Sheffield_Plain
Explore in full view
Which of the above visualisations would you prefer to use, when planning or deciding upon a cycling route? In the comment box, please explain why.
Choose one of the following answers
This question is mandatory
Arrow
Line-width
Line-colour
2D elevation profiles
3D elevation profiles
3D terrain models
Please enter your comment here:

Notice

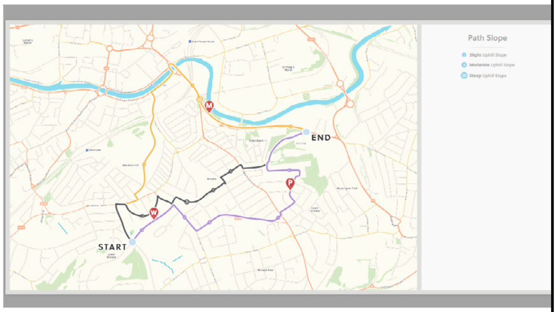
The following questions will ask you to select the **highest point** from a variety of options on the map. Please note the following:

- As soon as you have made your selection, press next to move onto the next question.
- All routes **start at the 'START' point and end at the 'FINISH' point.**

When you are ready to begin, please press next.

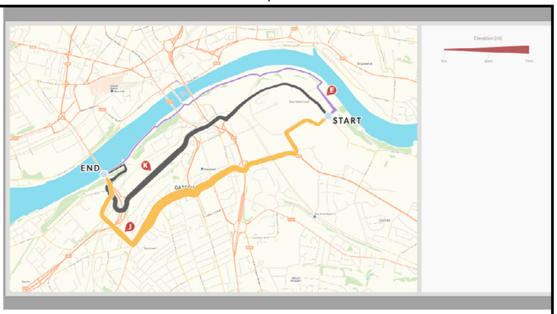
Point Height Questions

*
- Which of these points is highest?



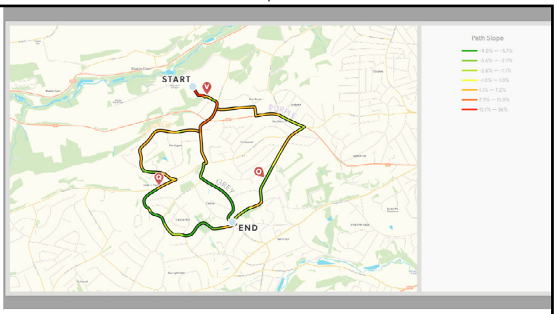
*
- Which of these points is highest?

- P
- W
- Don't know
- M



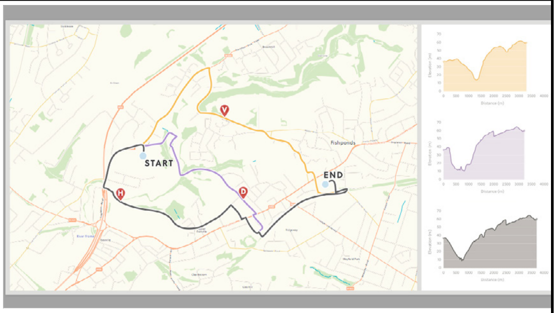
*
- Which of these points is highest?

- E
- Don't know
- K
- J



*
- Which of these points is highest?

- Don't know
- Q
- V
- O



*
- Which of these points is highest?

- H
- V
- Don't know
- D

3D_P_Alt_Newc_Height

*
- Which of these points is highest?

- T
- Don't know
- H
- D

3D_T_Alt_Sheff_Height

*
- Which of these points is highest?

- H
- Don't know
- W
- E

Notice

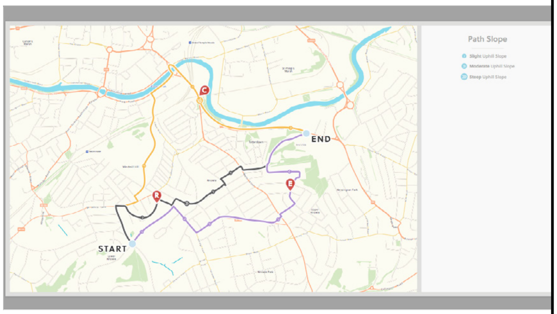
The following questions will ask you to select the **point which lies on the steepest slope**, from a variety of options on the map. Please note the following:

- As soon as you have made your selection, press next to move onto the next question.
- All routes **start at the 'START' point and end at the 'FINISH' point.**

When you are ready to begin, please press next.

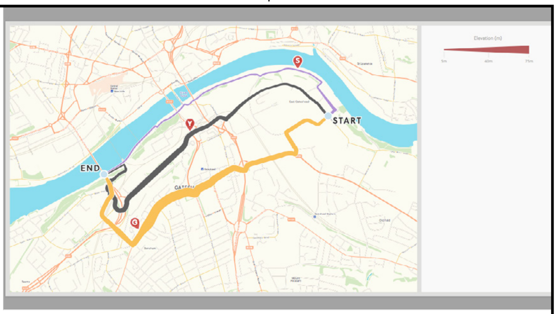
Slope Questions

*
- Which of these points lies on the steepest slope?



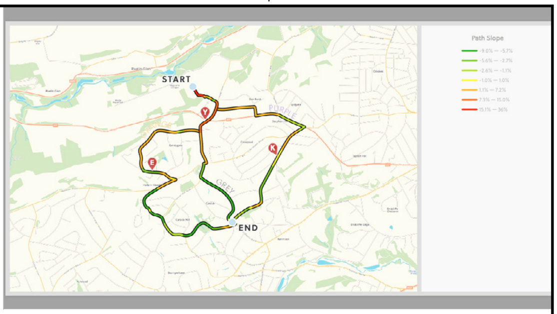
*
- Which of these points lies on the steepest slope?

- E
- R
- C
- Don't know



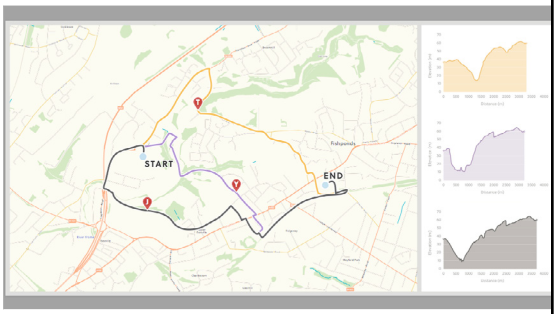
*
- Which of these points lies on the steepest slope?

- Don't know
- Y
- G
- S



*
- Which of these points lies on the steepest slope?

- K
- Y
- E
- Don't know



*
- Which of these points lies on the steepest slope?

- Y
- T
- Don't know
- J

3D_P_Alt_Newc_Slope

*
- Which of these points lies on the steepest slope?

- V
- D
- J
- Don't know

3D_T_Alt_Sheff_Slope

*
- Which of these points lies on the steepest slope?

- K
- U
- B
- Don't know

Notice

The following questions will ask you to select the **which of the three routes includes the most amount of climbing**. Please note the following:

- The 'most amount of climbing' is equivalent to the sum of the height of all ascents.
- As soon as you have made your selection, press next to move onto the next question.
- All routes **start at the 'START' point and end at the 'FINISH' point.**

When you are ready to begin, please press next.

Climb Questions

*

- Which of these three routes includes the highest total amount of climbing?

Note: the 'most amount of climbing' is equivalent to the sum of the height of all ascents.

3D_T_Alt_Sheff_Plain

*

- Which of these three routes includes the highest total amount of climbing?

Note: the 'most amount of climbing' is equivalent to the sum of the height of all ascents.

- Grey Route
- Orange Route
- Purple Route
- Don't know

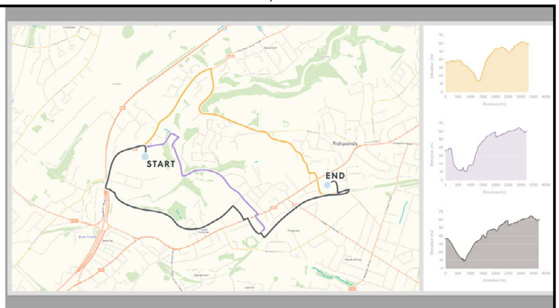
3D_P_Alt_Newc_Plain

*

- Which of these three routes includes the highest total amount of climbing?

Note: the 'most amount of climbing' is equivalent to the sum of the height of all ascents.

- Orange Route
- Don't know
- Grey Route
- Purple Route

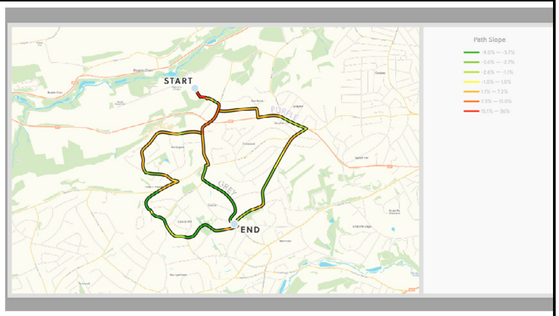


*

- Which of these three routes includes the highest total amount of climbing?

Note: the 'most amount of climbing' is equivalent to the sum of the height of all ascents.

- Orange Route
- Purple Route
- Grey Route
- Don't know

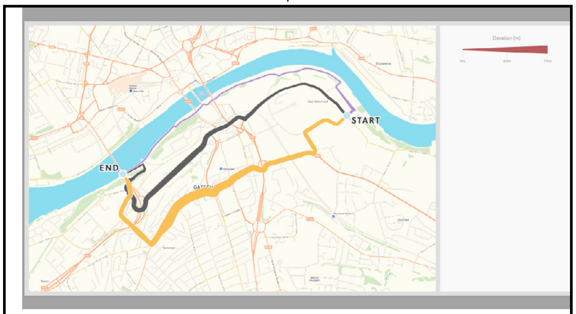


*

- Which of these three routes includes the highest total amount of climbing?

Note: the 'most amount of climbing' is equivalent to the sum of the height of all ascents.

- Grey Route
- Don't know
- Orange Route
- Purple Route

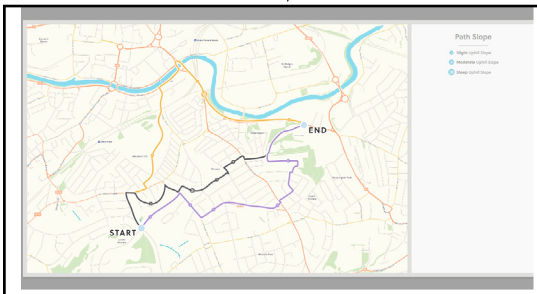


*

- Which of these three routes includes the highest total amount of climbing?

Note: the 'most amount of climbing' is equivalent to the sum of the height of all ascents.

- Orange Route
- Grey Route
- Purple Route
- Don't know



*

- Which of these three routes includes the highest total amount of climbing?

Note: the 'most amount of climbing' is equivalent to the sum of the height of all ascents.

- Orange Route
- Purple Route
- Grey Route
- Don't know

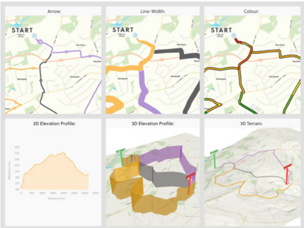
Notice

The following questions will ask you about your personal preferences regarding the visualisations shown. Please be honest!

When you are ready to begin, please press next.

Post-Preference

✱
When determining the **highest point**, which visualisation did you prefer?



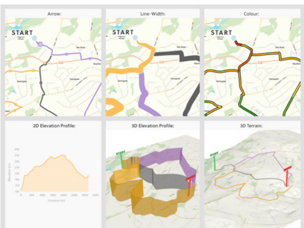
Please also give your reasoning!

❶ Choose one of the following answers

- Arrow
- Line
- Colour
- 2D Profile
- 3D Profile
- 3D Terrain

Please enter your comment here:

✱
When determining the **point with the steepest slope**, which visualisation did you prefer?



Please also give your reasoning!

❶ Choose one of the following answers

- Arrow
- Line
- Colour
- 2D Profile
- 3D Profile
- 3D Terrain

Please enter your comment here:

✱
When determining which of the three routes included the **greatest amount of climbing**, which visualisation did you prefer?

Please also give your reasoning!

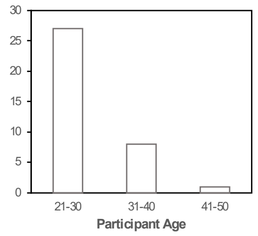
❶ Choose one of the following answers

- Arrow
- Line
- Colour
- 2D Profile
- 3D Profile
- 3D Terrain

Please enter your comment here:

2. <u>Demographic background of participants</u>

The following graphs show the age and gender makeup of the 36 participants who took part in the user study.



3. User preference comments

The following comments regarding visualisation preferences were provided by study participants before ('Pre-Preference') and after ('Height', 'Slope', and 'Climb'). Where no comments were provided, the preferred visualisation type has been removed from the list.

Pre-Preference	Comments
Line-colour	Using colour is very clear. I can see the steepness of each route immediately and compare them.
Line-colour	Most user friendly (meaning, you get all the info at one glance), together with the arrow system, yet you get more information about the terrain than just from the arrows.
Line-width	Here I can see which of the route has the highest elevations, although not precise, but it's possible to estimate the general pattern of the route steepness. Also I liked 2D-elevation profiles, but they look a bit frightening showing the highest elevation more steep (I think it seems so due to scale).
2D elevation profiles	I chose the 2D elevation profile because it shows you the distance of each route along with the elevation/climb required. It is also clean and very easy to understand versus the other options, which require a little more thought. My second choice would have been the arrows although that one makes no mention of distance. My third would be the colorful, line color version, but this one is almost too busy because you have to decipher the colors and then translate them into elevation profiles. In the end, you end up picking the one with the most green, but then again there is the lack of the distance factor.
Line-colour	It offers me the comprehensive detailed information about the route configuration at once, without too much zooming, scrolling, rotating and similar.
2D elevation profiles	The elevation profiles added on the side of the map are giving a helpful overview - I am able to know just from one look if the streets are flat or not. Second best in my opinion: Arrow
3D terrain models	The 3D terrain model shows the best how much of difference there is between the 3 routes. You can see easier where the hill-parts are and how many there are in general. you don`t really need to think about a percentage or something else explaining the map to you. You can actually see how the road is going to be.
Arrow	the 2D profiles gives a broad idea of how the terrain is and due to are in the same view is easy and faster take a decision. The arrow map is also good for me but needs more time to understand and look for every arrow. The line widthand 3D elevations are confusing for me; I couldn't identify changes easily. the line color makes thinks so much about compare colors between segments and 3D terrain model is cool but the landscape in general is not so steep so it also take time to decide which is the easy route.
Arrow	This visualization would be easier for me as a map reader to communicate to elevation of each route.
2D elevation profiles	It provides a visual and metric comparison of the routes in one section to the right of the map, which is easiest for me to read.
2D elevation profiles	Best overview and information given on point. Simple and good representation. Information is easy to depict at one glance.
2D elevation profiles	most realistic and I can explore the map according to my interest
Line-colour	Clear indication of up-/downhill directly in the map.
2D elevation profiles	I would prefer to use the 2D elevation profiles because, for me, it's the easiest visualization to understand the differences between the routes. I can better compare the distances and elevations of the three routes to see which route I would prefer to take.
2D elevation profiles	Gives the best optical impression how much/steep elevation you have in relation to the traveled distance.
3D terrain models	3D terrain model has the most directly visual effect of how the hilliness and steepness the path is to me.
2D elevation profiles	Because it provides me the total length of the paths and allows me to compare better between them. At the same time I can perceive better the terrain profile for every option and decide about the amount of effort I want to put in every segment of my trip. Depending on the next activity I am planning to do after the trip, I would decide for one or other option (if swearing more is not important for the next activity, I prefer the shortest path). The other options are interesting but doesn't allow me to get a complete sense about the alternatives. The second preferred would be the 3.Line-colour.
2D elevation profiles	By far the easiest to understand, without having to look around the map. The easiest to compare, too, especially if the service would allow to overlay the profiles with each other, to compare heights.
Line-colour	I like the profile view and line colour view but for me the line colour is better because I feel like the information is best represented here.
3D terrain models	I also like the other ones as they also depict the elevation quite nicely. Only the "arrow" map makes it hard for me to tell about the elevation. The 3D terrain model in my opinion is the one depicting the elevation the most "realistic" way perhaps.
Line-colour	I am used to reading these type of maps since they are similar to the Google Traffic Maps. It gives me a simple overview of the whole route.
Arrow	Seems like the most sleek design and offers a quick overview of the relative height differences. The focus of this visualization seems to be on the quick overview for the casual user as compared to giving specific degrees of slope which in my mind doesnt make much sense unless you are a super dedicated biker.
Arrow	arrow is the simplest and easiest way for me to digest the slope information. line wide: too bulky. line color: I might need that to inform me the traffic. line and profile : two steps information. 3D : I might not have time to digest the info when I am in a hurry.
3D terrain models	I found this one as the easiest way for me to receive all the information related with the terrain, and it doesn't have many figures and graphs which would make it not quite user friendly for the ones that dont have experience with map reading.
Line-colour	The map with line colours lets interpret the hills fastest without having to learn the functionalities of the map. It is very intuitive.
Line-colour	I used something similar before but for traffic situation, so this line colour visualization is more easier for me to understand and use colour to indicate the slope also gives me a really quick impression about overall data.

2D elevation profiles	I am familiar with the 2D elevation profile from using google maps for cycling, and I like it, especially when you can hover over a part on the elevation indication and this same point is indicated on the map. Great fun. Or at least... great utility. I also like the idea of the line colour model you gave above because it provides an indication that is intuitive and quick to understand upon seeing it for the first time; I could easily see the elevation change in each area and how it differentiated along the suggested route. I don't like the line width nor the 3d elevation models because they are not simple enough to immediately gain information about the elevation - it would take me a much longer time to understand what I was looking at as there is not a clear enough distinction along the route, if that makes sense. It looks more monolithic. The arrow map also is too subjective - what does a slight, moderate, and steep uphill mean? what about downhill - is this accounted for somehow? I would also want to know if there was a descent(s) enroute.
Line-colour	I like this one best, because it shows you the steepness of the cycle way instantly without having to scroll/rotate the map. Therefore it is easier to depict the easiest route. The colours are somehow self explaining (red = steep), so it is not really necessary to look at the legend only if I want to have more insight.
Line-colour	As it shows the elevation along with the extent of the elevation.
Line-colour	really easy to see how many hills there are in a route, with a single glance

Height Preference	Comments
2D Profile	Because the hight values are easily readable from the graph, although I couldn't locate 100% precisely the location of the point on the graph.
2D Profile	Again! And as well the 3D Terrain. The 2D again easy to tell with one look, faster! 3D great for more details, while tilting the map.
2D Profile	Clearest and fastest way to get the information. I don't like handling the 3D visualizations.
2D Profile	I found it easiest to identify the highest point when shown the 2D elevation profile.
2D Profile	I can see clearly and faster the highest point and compare with other possible alternatives without having to explore the visualization. (In cas I want to be efficient, if I have plenty of time I would prefer a 3D vis)
2D Profile	Easy to extract the highest point, because you can just read it from the graph, instead of comparing the lines on the map. It is the only visualization that actually shows numbers.
2D Profile	The 2 D profile puts exact numbers behind the visualization so in this respect I find it better, however it would have to be interactively linked with the map so that you can also see where that high point on the profile corresponds to the route
2D Profile	It clearly gives the elevation of each part of the route.
2D Profile	easiest to know the height of each point
2D Profile	it shows the terrain height and its more easy to understand the highest point
2D Profile	the 2d profile is the most distinct and precise, but only if you can relate a poitn on the map to an exact point on the elevation profile... line colour is not that useful since the elevation point on any part of the map is not known. i guess i could figure out if i looked at it for long enough what the relative elevation of one point to another was based on the starting location and then comparing the changing colour of each route, but this would be too time consuming and would probably not be accurate. the arrow model is way too vague and subjective and line width is not that useful either, in my opinion. the 3d terrain could be okay but would take time to think about to determine the highest point and still i probably wouldnt be sure
2D Profile	distinct height can be identified
3D Profile	Also 3D Terrain and Line were quite comfortable
3D Profile	I liked the 3D profile because you can use the pan tools to compare the routes and which one is the highest.
3D Profile	easiest to compare elevation between points
3D Profile	You can directly compare the heights of the points by using the interactive features.
3D Profile	With the 3D profile it is possible to determine the highest point very fast and easily. with the 2D elevation model I had to guess where the point might be and was not sure, because I am not very good in estimating. 3D terreain is also good but with this visualization I ALWAYS had to rotate the map first, which I would not want when I use an application. The first three visualization are equally good in my opinion, but the 3D Profile seems to be best for determining the highest point.
3D Profile	Would have been 2D profile, but hard to know where the point from the map is on the graph
3D Terrain	Also liked the line but there you might have a physiological limitation to seeing the exact width of the lines.
3D Terrain	Same as before - easier to compare the ropes.
3D Terrain	I prefer 3D due to is easy to analize the high and the location of the point. The 2D profile could be also very helpful if the location of the point would be easy to find in the profile; due to base map is a street map, was really difficult for me try to locate the point on the profile.
3D Terrain	I was more confident looking at the 3D terrain model to determine the highest point because it could be tilted to compare the elevations.
Line	I can see the elevation of the point directly from the width there.
Line	this visualization coded the value directly onto the map in an intuitive way, there was no need to look for a legend. It felt easy to spot the thickest line.
Line	Easiest to deduct and does not require additional actions like panning and/or zooming.

Slope Preference	Comments
2D Profile	as the height is clearly defined.
2D Profile	Because I can see clearly and faster the terrain profile.
2D Profile	but it would need an indicator of the position on the 2D profile. Then it would be perfect to determine the steepest slope.
2D Profile	Comparison is easier and handier, because you don't have to zoom in or tilt the map.

2D Profile	I liked the 2D Profile to determine the steepest slope. No fussing required with pan tools and such. It is blatant and obvious which is the steepest route.
2D Profile	Other representations are sometimes connected with to much visual comparison.
3D Profile	Also colour and 3D Terrain
3D Profile	In my opinion 3D profile is best for this task, as you can directly see the location of the point and its position in the elevation profile.
3D Profile	Same as before - would have chosen 2D profile, but it's difficult to guess where the point is on the profile
3D Profile	With the 3D profile it was easier to compare the ropes to figure out which one was the steepest rope. But I liked the 3D Terrain as well.
3D Terrain	Cannot explain exactly why
3D Terrain	I can see the slope directly from the 3D Terrain.
Arrow	It is the most simple visualization and I could just check which character is shown (one, two or three arrows). I didn't have to zoom and rotate the map or similar.
Arrow	The arrow presentation make the information easy to extract even at glance. By putting the arrows right into the slope breaks, it will very helpful for user to identify on which location they will face the "muscle burning" ones. 3 D presentation may provide interactivities and detail information, but it will take times to interpret the infos, and for me, I am not necessary need it for biking.
Colour	Even more direct information, because I don't have to switch between map and profile.
Colour	It gives the elevation profile of each section.
Colour	It is easy to find the steepest slope based on the colors.
Colour	none of them! but since i had to choose, i chose the line colour. the 2d elevation profile would be my preference if i could over on a point on the elevation profile or on the map and see that point mirrored in the other. the 3d terrain was either not hilly enough or the map was not not clear enough to distinguish that well between the elevation differences and slopes across the terrain. the colour one would be ok, if it had more distinct categories. sometimes the downhill slope category values were different from the uphill category values and to the lay user those might be annoying, confusing, or arbitrary numbers. the arrow one was only useful where the point in question was directly on an indicated hill, which was not the case in any of the questions. with the line width map it was most clear that it was a steep slope when there was an abrupt change in the line thickness, but often the change in thickness was continuous or too subtle.
Colour	The color literally shows the highest slope.
Colour	The color made it easiest to accurately find the steepest slope at a given point.
Colour	The colors specifically symbolize the degrees of slope, so I think that the color option is the best visualization for determining the point on the route with the steepest slope.
Colour	The steepness is directly colour coded, which made it very quick to read.
Colour	With this visualization you only have to look for the most red colour. All the other visualization require to compare the other routes or overlook the whole route the point was pinned in. The 3D profile would be as good, but here you need to rotate the map first to compare with the other points.
Colour	You can see really well visually where the steepest point is, the profile view is better for overall assessment of the slope profile but the colour is good for point assessment, so a combination of the two would be best
Line	Although I found the classes that were used and were differentiated by the colours not correct. The were not the same size and I did not see the logic in the classification.
Line	Its more easy understandable the terrain slope
Line	points between three elevations were easiest to compare without the need to further zoom in or rotate the map - 2nd best was 3D elevation: easy to compare but needed some extra exploration (zooming in rotating map)

Climb Preference	Comments
2D Profile	Again 2D Profile and Arrow Profile. But I think that with the routes beeing more similar the Arrow profile wouldn't be that great because it would be harder to tell the differences. That's why the 2D Profile is great, it is easy to compare and able to show the whole route and it's amount of climbing in one look
2D Profile	overall routes was easiest to compare by overall elevation profile no zooming needed
2D Profile	I couldn't identify the different routes in the color version.
2D Profile	I think the 2D elevation profile may have been my preferred option for determining the greatest amount of climbing, but the arrow visualization is also useful for me.
2D Profile	In the 2D profile the user can directly see the absolute height values of the route and thus determine easily, which route takes the highest amount of climbing.
2D Profile	Although it is tough to say where the climbs will be on the map, it is still the easiest way to see if there and how high the climbs are.
2D Profile	The 2D profile gives the best overall visual assessment of the route in a quick to determine manner
2D Profile	This profile gave values that could directly be interpreted, with the actual values, this made for an easy comparison across the whole route, but not for specific points.
2D Profile	easiest to see at a glance
2D Profile	Gives the best impression of the overall route. However, it is difficult to specify specific points on the curve. The 3D terrain would be more usable if the terrain would be characterized by more and steeper hills.
2D Profile	because i could compare elevation profiles and subtract the lowest point from the highest point
2D Profile	easy to see the climb in one go

3D Profile	After comparing all visualisation methods, I now prefer the 3D elevation profile method because it offers the most details about the climb heights as well as the distance (curvy or not) from an intuitive standpoint. By manipulating the pan and looking at the 3D terrain plus the transparent ribbon height, you gain the best sense of how long you will be sweating on your ride. :)
3D Terrain	Because it was most understandable from all the visualizations. I could explore the terrain and then make the conclusion on my personal feeling of the amount of climbing.
3D Terrain	Its more distinct and shows better the terrain
Arrow	For me determine this parameter was difficult due to is related with the high
Arrow	easiest and simplest way to obtain the information.
Colour	The colours make it obvious to me that where is uphill and where is downhill. I can calculate it based on that.
Colour	Easy to interpret.
Colour	Easy to see it due to colours
Colour	The colour-map gave me a better overview and I didn`t have to look that close to see where the most amount of climbing was. I did like the 2D elevation Profile as well in that case.
Colour	The breakdown of more than three slopes using color helped to compare the slope and the distance.
Colour	Because I can get a faster impression about the effort required, adding reddish segments.
Colour	i like the three colour-thing, because its pretty clear to understand
Colour	It says the profile at each section and at each place along the route.
Colour	colour, as the number of transitions are shown by the fluctuations in colour.
Line	Though very vague it is the fastest for the visual system.
Line	the 3D visualizations required to be rotated first to get an idea of the uphill slopes. The 2D visualiations fit better. The line width gives the best overview of uphill slopes because when looking at the maps with arrow and colour I had to estimate how many times a bigger uphill slope was visualized and as I said before, I'm not good in estimating