Exhibit Prototype Testing
Lessons from the Launchpad Project

Sofie Davis

Audience Research and Advocacy Group
Science Museum; Learning

Lessons from Launchpad – Prototype Testing
Science Museum: Learning
Audience Research and Advocacy Group
Sofie Davis / December 2007

Audience Research & the redevelopment of Launchpad

Launchpad is an interactive gallery for 8-14 year olds. Audience research has played a vital role in ensuring the redeveloped gallery is engaging, accessible, and educational to its target audience.
In the gallery, mechanical interactive exhibits demonstrate core physics phenomena such as light, electricity and forces, and aims to encourage scientific thinking such as questioning the world around you.

Audience Research was conducted throughout the Launchpad redevelopment project, and included a range of different methods, including focus groups, usability testing, interviews and observations.

This document will summarise the *formative* stages of evaluation, where prototypes of the exhibits were built specifically to be tested with the target audience. Formative testing was designed to identify potential barriers to effective visitor interactions, but it also taught valuable lessons about how to prototype test exhibits and highlighted some key features necessary when producing a successful interactive exhibit.

Lessons Learnt about HOW to prototype test exhibits

1. It is essential to have clear aims and objectives for prototype testing. Before testing an interactive exhibit, ask the team to define its learning outcomes and describe what they would like visitors to get from the experience. As an evaluator you must consider:
    - What are the most important things you are trying to find out?
    - How will this information help you improve the exhibit?

2. When assessing prototype exhibits, three central areas must always addressed:
    - **Motivation** – Do visitors want to use it? Do they enjoy using it?
    - **Usability** – Can visitors work out how to use it? Do they know what to do with it?
    - **Content** – Do visitors understand what the exhibit is about? Do they recognise aspects of what it is trying to show them?

    During the evaluation of prototypes always look out for motivational, ergonomic and intellectual barriers – it is these barriers that you are seeking to identify and remove in order to improve the exhibit for the target audience.

3. Prioritise the exhibits to be tested. Due to time constraints it wasn't possible to test three prototypes for each of the 50 exhibits. Therefore, it was important to identify which exhibits had to be tested and which exhibits would benefit from the most testing. For example, those exhibits which were completely new, had substantial changes made to existing versions or which we knew had significant problems (motivational, ergonomic or intellectual barriers) were prioritised for testing. Those which had already been extensively tested (e.g. the APE exhibits from the Exploratorium) or which there was extremely limited scope (e.g. Air Cannon) for change were not tested.

4. Always plan to test more than one prototype for each exhibit. Ideally, three prototypes for every exhibit should be planned for. Testing several versions of the same exhibit ensures that recommended changes are successful and reduces the likelihood of introducing new problems with new versions of the exhibit.
    - When re-testing adapted prototypes look specifically at how the *changes* are affecting visitor behaviour, rather than gathering more detail of what was already known.

5. It is extremely important to find a suitable area where testing can take place. We used an old gallery space (Things) for cued testing. Key qualities for a testing area include:
    - ❖ Readily available: often you will have to do additional testing at short notice and in this scenario do not want to be negotiating with other museum teams for use of a space.
    - ❖ Electricity supply – many exhibits have an electrical component.
    - ❖ No-public access – it is important that you can leave prototype exhibits out without risk to the public otherwise you will have to spend large amounts of staff time guarding the exhibits.
    - ❖ Close to where you are going to recruit members of the public – this will make it easier for you to recruit and bring visitors to and from the space.
    - ❖ Reasonably quiet – so that you hear visitors responses both during their use of the exhibit and in response to your questions.

    If you plan to do un-cued testing on gallery, it is important that this is conducted in an area which attracts a similar target audience to the one that the exhibit is aimed at. There is no point testing an exhibit aimed at 8-14 yr olds with adults.

6. There are advantages and disadvantages to using both cued and un-cued testing:
    - ❖ Cued testing is when you recruit visitors to test with i.e. they have been briefed that they are going to be using a prototype.
    - ❖ Un-cued testing is when you don't actively recruit users, but instead put a prototype out on gallery and allow visitors to come across it and use it in a non-briefed, naturalistic setting. w

	Benefits	**Problems**
Cued Testing	▪ Is safer and can be better controlled if a prototype is still a rough mock-up. ▪ Visitors focus on the prototype alone, providing data on how they interact with and react to *that* specific exhibit. ▪ Visitors tend to stay with the exhibit for longer so more types of interactions can be observed. ▪ Visitors have agreed to be interviewed and these can therefore be longer and more detailed. ▪ It is possible to leave the exhibit there if the set-up is complicated and / or testing carries over a few days.	▪ Visitors behave differently when they know they are being watched. They tend to put more effort into the interaction and an exhibit might therefore appear more successful than it actually would be out on gallery. ▪ Timing visitors is not a useful way to measure engagement in this setting – they tend to stay for longer when watched. ▪ During cued testing more visitors read labels and therefore have seen instructions they may pass over in a gallery setting. ▪ Recruiting suitable visitors to test on quiet days can be difficult and delay results. ▪ The availability of the room affects the testing timetable.
Un-Cued Testing	▪ Shows how visitors would behave naturally in the gallery setting. For example, they may not attentively read labels or persevere with an exhibit they do not understand or find boring. ▪ Gives a more accurate idea of how long visitors are prepared to use the exhibit. During Launchpad cued testing this varied from 3 seconds for a poor interaction to 26 minutes for a very good interaction. ▪ Un-cued testing can generate more data, more quickly than cued testing since	▪ The exhibit needs to be made safe enough to be placed in a public area, e.g. no cables exposed, no finger traps, sharp edges or unstable surfaces. ▪ Testing needs to be done on a day when the right age range is in the gallery (often they were younger in the old Launchpad than we aimed at for the new). ▪ If interviews are given after un-cued observation, these need to be much shorter as visitors are unprepared and easily distracted by wanting to explore the gallery. ▪ There is less flexibility over the environment

	there is no recruiting time and often interactions are shorter and continuous. ▪ Explainers working in the gallery could get an idea of the testing process and see the new exhibits being introduced.	and positioning of the exhibit - which must not block any walkways or other exhibits. ▪ Finding suitable tables and electrical outlets can be problematic. ▪ The gallery team leaders always have to be consulted and asked for permission to ensure set up is Ok in the gallery. This uses their valuable time.

7. Regardless of whether you are using cued or un-cued testing, all prototypes need to be risk assessed prior to use with the public. It is important that a number of staff are trained in risk assessment prior to the testing period.

8. Prototype testing involves working with many teams, for example the project team, the gallery staff and the workshops and designers making the prototypes. Informing all those involved of what is being done, giving them good notice of when their help is necessary, and showing some appreciation makes the work a lot easier for everyone.

9. When designing an observation sheet match what you are looking for to the learning outcomes of the exhibit and aims of the testing. What might you look for as evidence of successful interaction? What would show that users were not engaged?

10. In addition, it is important to look for behaviours that you would like to see/expect to see but which are absent. These behaviours are often missed in reporting but give valuable information about the success (or otherwise) of the exhibit for the audience.

11. Ensure that you have enough data from the interviews to be able to support your argument, but don't be excessive in gathering the same data over and over. During Launchpad testing we found that for cued testing, somewhere between 10 and 20 observations and interviews produced sufficient data to highlight the range of depth of barriers to use and engagement with an exhibit. Typically this would take 3-5 days to collect. During un-cued testing on gallery, we found that that 1-2 hours of observation in a reasonably busy gallery could give a good idea of how visitors behave with an exhibit.

12. Interview questions need to be specifically tailored to the exhibit and questions you most want answered. What will the questions tell you about: what visitors thought the exhibit was about, whether they enjoyed using it, or what they thought it was about. Consider using drawings, pictures and prompts to help understand visitors' thoughts.

13. Document everything well while it is still set up:
    - ❖ Testing proved that the spatial relationship between different elements of an exhibit is important in guiding visitor behaviour. By recording these measurements clearly, better advice can be given about the final set up in the gallery.
    - ❖ Taking pictures of how an exhibit was set up, or the different actions of visitors (for example how they placed blocks) can help the team understand more clearly how an exhibit was used and provide evidence to support recommended changes. They also act as a good reminder months on of what the prototype was like.

14. Invite the project team to watch visitors using the exhibits. This helps them to understand the kinds of interactions you describe and where the exhibits can be improved. Do this with caution however, people will often see what they want to and the interaction they watch may not necessarily be representative of how most visitors respond.

15. Ask other members of the Audience Research team to watch the testing - their input and experience can be really useful. It is also useful to discuss findings with the Audience Research team before feeding back to the project team to ensure findings are clear and focused on realistic and useful recommendations.

16. Provide feedback quickly to the project team. Writing up the findings always takes at least as long as the time spent collecting the data. If the testing schedule is tight try to provide the team with top line findings and key recommendations so they can progress without waiting for the full report.

17. Provide feedback in more than one way – project teams often don't have the time to read reports. Feeding back face-to-face provides the project team with a valuable opportunity to ask questions and for you to ensure that the implications of the findings are understood.

18. When interpreting testing data, focus on the underlying issues affecting visitors' interaction. If an exhibit is failing there can be many things that need to be improved, but look at what the core problems are. For example common underlying problems that we found included:
    - the exhibit not having a clear challenge;
    - the activity having too many options so the visitor gets overwhelmed;
    - the interaction not having a clear starting point that indicates what to do and why;
    - the exhibit effect is underwhelming to visitors.

    Changing the shape of a button or the text of a label will not solve any of the above problems. Instead the underlying messages, activities and responses of the exhibit had to be challenged.

19. When writing up findings from testing, highlight the most important recommendations and key findings at the top of the report where they make the most impression. Prioritising issues and being concise make it more likely that the team will read the report and respond to the recommendations. Including pictures and quotes from visitors can also help to build a stronger picture of visitor responses.

20. Some good news goes a long way. If there is a nice quote or example of success, let the team know so you are not always the bearer of negative feedback.

Lessons Learnt about WHAT makes a successful interactive exhibit

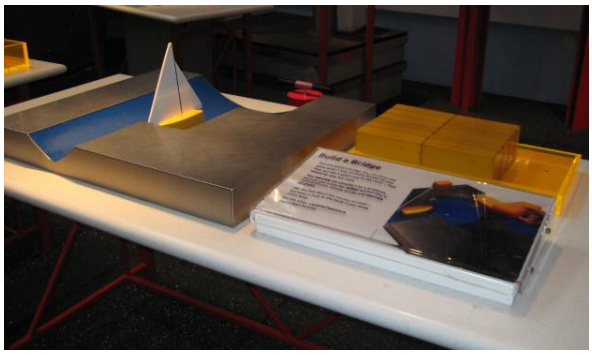
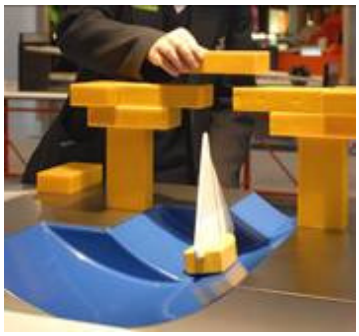
There is no single recipe for making a successful exhibit. Each exhibit is unique and will have its own issues affecting usability, enjoyment and learning. However, through prototype testing it is possible to identify some of the key qualities that help to make an exhibit interesting, challenging and engaging.

1. A successful interactive is one where barriers preventing visitors from engaging have been identified and removed. Barriers can be:
    - **Physical** - visitors are not able to use or understand the exhibit design / interface / controls.
    - **Emotional** - visitors are not inspired to use the exhibit, they do not feel confident, delighted, curious or in control.
    - **Intellectual** – visitors are not able to access the content the exhibit is about, they do not understand key messages, or do not feel like the exhibit is for them.

2. Ideally a mechanical exhibit needs to have a clear starting point, but develop through to a more open end point. Visitors need to be constrained in what their initial interaction is so they are not overwhelmed with choices or confused about what to do first. As the interaction progresses, there

should be opportunities to extend the activity, so there is open exploration rather than a single solution. This kind of interaction can occur in many different types of exhibits. For example:

- Build a Bridge – the design and instructions on this exhibit make it obvious what the visitor is expected to do, but there are 3 levels and many ways of completing it, most of which involve experimenting with counterbalance.

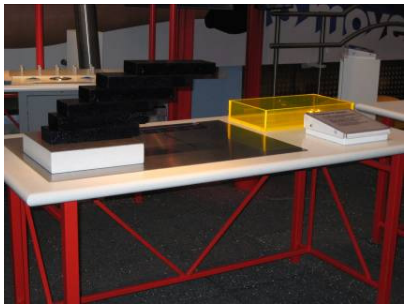



Build Bridge exhibit with 3 levels of difficulty shown by the 3 differing widths of river.

- Thermal Imaging Camera – this exhibit gives instant feedback showing that you are looking at an image of yourself on a screen. The layout and available props then encourage visitors to realise the camera picks up heat and experiment openly with how heat is transferred.


The Thermal Imaging Camera Exhibit with screen and props.

3. There are different qualities that will engage and motivate visitors to use an exhibit, and each exhibit should have at least one if not a combination of the following:
   - Encourage **open-ended** exploration where the visitor in is control, can make choices and hypothesise about of their own interaction. E.g. Thermal Imaging Camera
   - Offer a **challenge** where visitors are enthused to achieve something and want to feel pride in their success. E.g. Hang Over Problem
   - Be **surprising** or counter-intuitive so visitors are intrigued or curious to find out more. E.g. Sound Bite or Seeing Through Walls (periscope).
   - Be a visually **beautiful** demonstration of scientific phenomenon. E.g. Icy bodies.



The 'Hang Over Problem' exhibit

The 'Seeing Through Walls' exhibit

The Icy Bodies exhibit showing dry ice sublimation

4. It's all about **me**! Visitors have a tendency to relate everything to themselves and enjoy using exhibits that can tell them something about themselves. Building in challenges and relating exhibits to life can therefore help to engage people, but there is also a danger visitors will focus on themselves and be distracted from the intended exhibit content:
    - When an exhibit offers a challenge, visitors will often assume that they are testing their own abilities.
    - Visitors relate well to exhibits which they can link with aspects of familiar everyday life.
    - Even when visitors are investigating abstract physics phenomena, they do so with their own abilities in mind. For example, with the Big Machine pulley's pod, lifting weights with different pulley systems was perceived as a test of personal strength. With Sound Bubbles, visitors saw a real-time image of themselves on a large screen and their dancing and jumping around in front of this distracted them from the activity, and the exhibit had to be left off gallery.

5. Visitors interpret exhibits and labels very **literally**:
    - Visitors will often follow instructions without thinking through what the result might be. For example, the Thermal Imaging Camera exhibit asked visitors to lick and blow on their hand (which would then show up cold on the screen). Visitors would read the label and lick and blow on their hand expecting something to happen without looking at the screen. The instruction to 'look at the screen' had to be added.
    - The use of props or metaphors must be done with caution, be tested if possible, and be clearly related to the subject; otherwise they become a diversion from the aim of the activity. For example, the exhibit, 'Grab the Bling' was originally tested using a parabolic mirror to reflect a duck with a £5 note in its bill. The instruction to 'grab the bill' was not understood by visitors (some visitors didn't know that 'bill' referred to a note as well as a beak) and some children interpreted this as an exhibit to teach you not to steal.

6. However, even though visitors take things very literally, often they do not understand clues that *you* may think are completely obvious. Don't make **assumptions** about how visitors will behave or what they will understand.

- For example; when testing Build a Bridge, to try to encourage visitors to place the large red blocks vertically instead of horizontally and build over something, a red rectangle shape the same size as the base of the block was placed at three points along a 'river' made of blue paper. Visitors totally ignored the red rectangles and sat the blocks in any order in the 'river' itself - totally missing the core content and challenge of the exhibit.

7. The **positioning** of an exhibit *and* of the different elements of an exhibit in relation to one another, are really important in guiding visitors' interaction:
   - The spacing between components of an exhibit affects visitors' understanding of how these elements are related. For example, The Thermal Imaging Camera has a screen, a bench with props on and a background panel. Testing showed that if the bench was too far from the screen, visitors would ignore the screen and simple try to investigate the props. The layout needed to be close enough to show a clear relationship between elements, but provide enough space to allow people to recognise a clear area where interaction was supposed to take place – almost like a small stage around which other visitors would have to watch rather than walk through and disturb.
   - The placing of an exhibit in the gallery also affects how visitors will use it. For example, when Downhill Racer was placed with the top end of the sloping rails against a wall, visitors gathered at the lower end and tried to forcefully push weights *up* to the top. When placed parallel to a wall, with a label emphasising the top end of the exhibit, many more visitors used the exhibit correctly by letting the weights roll *down* the rails.

8. **Interference** between visitors, where one visitor is able to disrupt the interaction of another, needs to be avoided in the design of the exhibit.
   - For example; the first prototype of Noises in Disguise (Hip Sounds) was designed to be multi-user, but it allowed visitors to press buttons that would affect the feedback other visitors heard. This led to a confusing experience where users were unclear about what they were affecting and they therefore lost control over their interaction.

The first Noise in Disguise prototype

9. However, preventing interference should not reduce visitors' opportunity to **communicate** with each other at an exhibit. Encouraging good communication at an exhibit helps to create a successful experience for visitors, since they are able to communicate their ideas and engage in social learning. For example;

- The Light Table was designed in the shape of a clover flower to allow visitors to have their own interaction area where they would not impact on the space of other visitors, but kept a feeling of intimacy and an openness of spacing that meant visitors face each other and can clearly see what the other is doing.
- The Periscopes exhibit involves visitors looking through two ends of a periscope and seeing each other despite a wall standing between them. It works best as a 2 person activity, and therefore this wall needed to be large enough to obviously block the line of sight through the eyepieces, but small enough to allow visitors to peer round it to communicate about what they could see.

10. It is important to design exhibits and labels that encourage adults to **scaffold** children's interactions – this relationship supports and improves the learning experience of both. Methods by which to encourage adult participation with children's learning were introduced in the Launchpad gallery:
    - When interviewed many parents felt they were not confident enough in their scientific knowledge to successfully engage with their children's interactions. If labels are short and clearly highlight key words and phrases, parents feel more supported and are more able to feed the information on to children.
    - At four exhibits simple video labels were introduced, showing the exhibits being used by adults and children and making the goal of the exhibit more apparent. These have prompted more adults to help support their children's interactions

11. While it is important for users to feel confident in their ability to complete a challenging exhibit, don't be too afraid to have a challenge that is **difficult**. If a visitor is engaged in the challenge, they will spend the time pushing their interaction further. For example:
    - Build a Bridge has three widths a bridge can cross, the narrowest being the easiest and the widest being hard. When there were only 10 blocks available this last width was a tough challenge, but during testing some children spent up to 13 minutes completing it.
    - Circuits exhibit has different levels of circuits to complete, from simple to quite complicated. For those who want to use or build on their knowledge these harder circuits give them something to work to.

12. Interactives can look **sophisticated**. Simply because an interactive is meant to be robust and used by children, does not mean that it has to be made of bright primary coloured plastic. Launchpad is designed for 8 – 14 year olds, and therefore needed to be seen by this group as for them. This was particularly hard for the older end of the age range who differ greatly in their levels of knowledge, needs and expectations. By producing exhibits that look like they are elegant or delicate can lead to visitors treating them with greater respect and care than those designed to look robust enough to survive rough behaviour. For example; the final tapper boxes used for Vibration Station kept their delicate feel being reasonably small, white in colour with coloured lighting inside, and light to hold. In addition, in order to encourage adults to participate to support children's learning it is important that the exhibit doesn't appear childish.

13. Providing for increased **accessibility** for those with disabilities, does not mean dumbing down or simplifying the content of an exhibit. Often changes to make an exhibit more accessible lead to less barriers for every visitor. Ensuring good accessibility enforces good design principles. For example:
    - In Sound Patterns, where silicone oil moves in patterns as sound vibrations run through it, the clear silicone was replaced with blue silicone to improve its contrast against the background.

This has led to the feedback being more obvious and it being a more visually beautiful exhibit for all.

- For Vibration Station, the start/stop mechanism had to be usable for those with manual dexterity problems, and is therefore more straight forward for all visitors. There is also a range of feedback from this exhibit; different colour lighting indicate whether it is in record or play mode, and as the exhibit works there is both sound and vibrations.

14. How visitors understand **what to do** with an exhibit will depend on the clues we provide for them. These clues need to be both physical (in the exhibits design) and interpretational (in its labelling). It is necessary to account for different learning styles when designing an exhibit and its labelling, with some visitors finding it easier to read instructions, and others finding explanations accessible through direct trial.

15. The way an exhibit is **designed** lets visitors know how to use it. The expected actions need to feel intuitive and obvious. The relationship between elements, the size of elements, the shape of buttons or handles will all indicate what the interaction should be and where the most important aspects of it are. Ideally visitors should be able to use an exhibit even if they don't read the label (as so often they don't or can't). However, as described in point 2, this design should not be so rigid as to prevent visitors from extending their interaction once they understand what to do. For example;
    - The first prototype of Noises in Disguise (Hip Sounds) had 4 equally sized buttons, a blank surface where the interaction was supposed to occur and one large and two smaller screens. Visitors were confused about what they were supposed to do, pressed the buttons in a random order, and paid more attention to the screens which gave confusing feedback because they were not able to use the main interface area. The final much improved version has one large feedback screen above the clear interaction area, and one button to the side that visitors can chose to use but can see is not essential.
    - Another example is Yacht Racer where visitors change the direction of a sail to move a boat across a table with fans. The boats are on wheels, but when these wheels were exposed visitors were more likely to push the boats. When the wheels were hidden, visitors treated them more like boats and pushed less, relying instead on the wind and therefore engaging with what the exhibit wanted to show.

16. Exhibits often do not work if they rely on visitors completing a **sequence** of actions to get to the end result. A sequence of steps is unlikely to be completed by visitors and if an exhibit hinges on any strict order of actions being followed, the interaction will fail. Ideally interactives should be asking visitors to do simple procedures that then lead to complex outcomes.

17. The **feedback** a visitor receives from an exhibit must be almost instant. When a visitor interacts with an exhibit something needs to happen almost at once, if it does not a common reaction is to assume that the exhibit is broken and/or the visitor is quickly distracted and walks away.

18. The interpretational **labelling** of the exhibit is also essential in helping visitors understand what to do with an exhibit and what the exhibit is about. Having video labels, picture labels and 'element text' all helped visitors better understand the Launchpad exhibits
    - 'Element text' were instructional words placed on the exhibits themselves, so visitors would immediately see them and be given a clue about what was happening. For example; Downhill Racer had 'start' and 'finish' added to the two ends of the exhibit to show the direction of the

wheels. Surprising Shadows had '4 lights' and '8 lights' added next to buttons to change the four to the eight lights. This highlights to visitors that the number of lights is an important feature.

- Picture labels were added to four exhibits where visitors were having difficulties understanding what to do. They provide an instant starting point for visitors to see how to begin their interaction. For example; with Build a Bridge the picture label shows that visitors should place the blocks vertically, outside of the river, and other blocks can be added on top to balance across the gap. The picture is incomplete so visitors are not directly given a solution.
- Video labels are short videos showing children and adults using an exhibit correctly. They give instant information about how visitors should behave and what the exhibit can do, conveying complex actions that would be difficult to describe in text and would be unlikely to be read.

19. If an exhibit relies on a **reset mechanism** to get it back to the starting point of the interaction, this needs to be either clearly labelled, or if part of the design, subtle enough not to detract from the actual activity. For example;
    - In the video label for Arch Bridge the bridge is shown being dismantled first, before it is rebuilt. Visitors on gallery may approach a completed bridge and therefore need the label to indicate that the starting point of the activity is with the bridge dismantled.
    - For the Archimedes Screw Big Machine Pod, the ball bearings are lifted up to the top of the screw, but then need to return to the bottom to let the process be repeated. The first prototype of this had the screw on a lever which tilted so the top became the bottom and vice versa. When tested visitors were distracted by rapidly tilting the exhibit up and down so the balls moved without the use of the Archimedes screw. The solution was to keep the screw stationery and add a small tube the balls would roll down from top to bottom. This did not distract from the main action and added the benefit of giving the journey of the balls through the screw and back down a more pleasant sense of purpose.

20. The title and design of an exhibit help visitors to understand **what the exhibit is about**. The visual clues emphasised in the design of the exhibit, and the way it is labelled, contribute significantly to visitors' understanding of what they are doing and seeing. For example;
    - When periscopes was tested with very obvious light sources shining on the eyepieces, many more visitors said that this exhibit was about light and needed light to work.
    - An exhibit title can help visitors either to realise what the exhibit it, or what they are supposed to do with it. Abstract titles that rely on slang or a pun may seem more light hearted, but cause more confusion. For example; 'pulleys' and 'wheel and axel' for the Big Machine Pods instead of 'Pulling Power' or 'A Wheel Advantage'. The first titles are less confusing and introduce science vocabulary. Another example is 'Build A Bridge' instead of 'River Bridge'. This title describes simply what the visitor should do.