Central European Multimedia and Virtual Reality Conference (2005)
J. Zara, J. Sloup (Editors)

Menu Selection in Desktop Virtual Reality

Jurriaan D. Mulder

Center for Mathematics and Computer Science, CWI
Amsterdam, the Netherlands
mullie@cwi.nl

Abstract
Menu systems are a standard interface technique in 2D desktop systems. Using menu systems in virtual reality systems however, is often problematic. In this paper, we present five techniques for menu selection in a 3D desktop virtual reality system: a 3D ray-casting technique, two projective techniques, one technique based on 3D positioning, and one technique based on the standard 2D desktop mouse. The different techniques are compared in a user study. Furthermore, 3D menu location is also considered in the evaluation. The results show that menu selection can very well be performed with 3D devices, provided that the 3D movement is constrained.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality; H.5.2 [Information Interfaces and Representation]: User Interfaces—Evaluation/methodology, Input devices and strategies.

---

1. Introduction

Menus provide a powerful means of human computer interaction. The Windows, Icons, Menus, and Pointer (WIMP) metaphor is widely used as the standard interface in 2D human computer interaction. In virtual reality however, no standard has yet been established for the use of menus in 3D.

Over the past number of years, there is an increased interest toward more affordable environments for virtual and augmented reality (VR/AR). Many types of 3D applications do not require CAVE-like or Head Mounted Display (HMD) based, fully immersive environments. Often, *fish tank*, *desktop*, or *dexterous* types of environments are sufficient [Dee92, PS96, WF96]. In principle, these type of environments can still make use of the standard WIMP metaphor for system control and other types of interaction. However, this is not always desirable. When working in a 3D application using 3D interaction devices, a context switch and an input device swap is required in order to be able to use the standard 2D WIMP metaphor. By providing a menu manipulation technique based on a 3D input device, such a switch can be avoided.

Another aspect that comes into play when using menus in a 3D environment is that of menu position. Whereas on a 2D display menus have a position that is 2D or 2.5D (in case of overlap with other windows), in 3D displays menus have a 3D position, that is, they have an apparent depth. The depth position of a menu might be of influence when manipulating the menu.

At the Center for Mathematics and Computer Science, an affordable desktop environment for near-field VR/AR is being developed; the Personal Space Station (PSS) [ML02]. The emphasis of the PSS is on direct 3D interaction, ergonomics, and low costs. The design and a prototype of the PSS are depicted in Figure 1. The PSS is a mirror-based colocation display system built out of low-cost, off the shelf components. The user is seated in front of a mirror which reflects the stereoscopic images of the virtual world as displayed by the monitor onto a *Virtual Focus Plane* (VFP). Three-dimensional interaction is performed using optically tracked, tangible 3D interaction devices. The user reaches under the mirror into the virtual world without obscuring the image or colliding with the monitor. Therefore, interaction is performed directly in the 3D workspace.

During the development and application of the PSS, the problem of 3D menu manipulation for user input was encountered. Although it is possible to use a standard desktop mouse in the PSS, it is often undesirable and disturbing having to switch interaction devices. Therefore, it should be

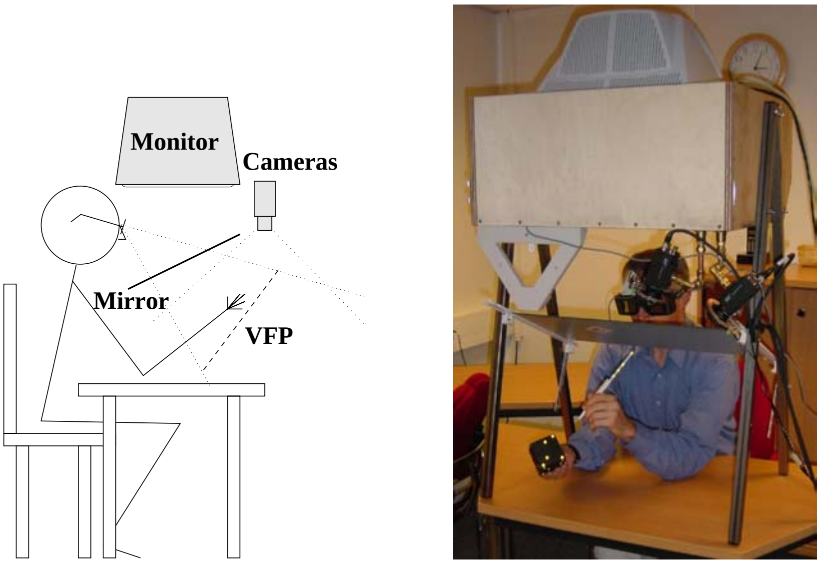


Figure 1: *Schematic side view and a prototype of the Personal Space Station.*

possible to manipulate a menu using a 3D interaction device. Furthermore, it should be possible to position menus at various locations inside the 3D workspace, including different depths. This can be helpful to prevent menus from obstructing the view on the 3D objects of interest and vice versa, or to position them close to a particular object the menu operation will apply to.

In this paper, we present a comparison of a number of menu selection methods for desktop and dexterous virtual environments. We compare four methods that use a 3D interaction device and one method that uses the standard 2D desktop mouse. Furthermore, we consider the location of the 3D menu in the workspace: at, in front, or behind the focus plane. In the next section, we will briefly review related work on menu systems for virtual environments. In Section 3, we discuss the different menu selection methods we considered in this study. In Section 4, a user study is described, and in Section 4.2 the results of the evaluation are presented. Finally, in Section 6 the conclusions and areas for future work are presented.

2. Related Work

Many types of menu systems have been used in 2D desktop systems, such as pull-down, pop-up, and pie menus. Most of these have been ported for use in virtual environments, but there have also been some new paradigms for menu systems in VR. Jacoby et al. describe the use of pull-down menus in VR [JE92]. Selection on these ‘floating menus’ can be achieved by ray casting or image plane occlusion techniques [PFC*97]. The position of such floating menus in 3D space is static, or they can be fixed relative to the users head [BH95] or the users body [MBS97].

A different approach is to use two handed interaction in combination with tangible input devices, such as a pen and tablet [SES99, LSH99]. It has been shown that the use of such devices and the tactile feedback they provide during interaction can significantly improve user performance [KL04].

While menu systems and menu selection has been extensively studied in standard 2D desktop systems, the number of formal user studies on manipulating menus in virtual environments is limited. Bowman et al. present the design of the TULIP menu system (based on Pinch Gloves) and compare their system to floating menus and pen and tablet menus in an empirical evaluation [BW01]. These menu systems and evaluations are performed in fully immersive virtual environments based on using HMDs.

Our study is aimed at using pull-down menus in desktop and dexterous VR. There have been empirical studies on user performance in 3D interaction for these type of systems, e.g. [AW00, ABW93, WAB93, WR99], and new interface techniques have been presented, e.g [SPHC95]. However, to our knowledge there is no study available on single-handed menu selection in these type of environments.

3. Menu Selection Methods

We have designed and implemented five different menu selection methods. One method uses the standard 2D desktop mouse, the other four methods use a 3D input device. These are a standard ray-casting technique, two projective

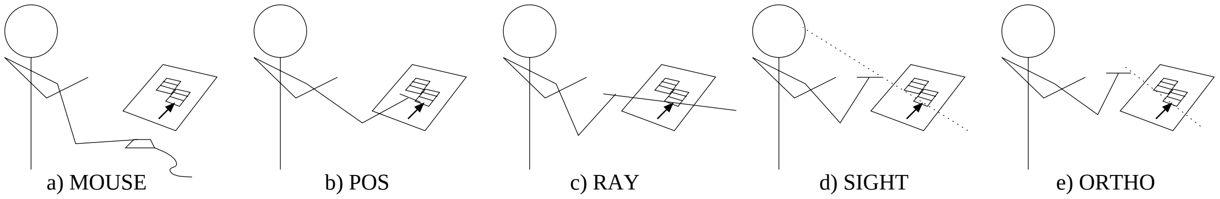


Figure 2: *3D Menu selection methods.*

techniques (orthogonal and line of sight), and a 3D position technique.

3.1. 2D Desktop Mouse

In most desktop virtual environments, the standard 2D desktop mouse is still available for interaction purposes. We can extend the use of the mouse to 3D menu selection by simply transferring mouse motions to a 3D cursor whose movement is constrained to the 2D menu plane in 3D space, see Figure 2a. We will refer to this method as the MOUSE method. This approach is expected to be easy to use, since it is very similar to the standard 2D menus. In addition, there are no constraints as to where the menu is positioned. It does however, require an input device swap when working in 3D.

3.2. 3D Position

In this technique, the 3D cursor position is directly coupled to the position of the 3D input device. Hence, the 3D input device is to be positioned at or near the menu plane to perform a menu selection, see Figure 2b. We will refer to this method as the POS method. The POS method is only applicable to menus located within reach of the user.

3.3. Ray Casting

Ray casting is a well known technique in virtual reality. It has been used for menu selection in many applications and environments. Generally, a 3D ray is extended from the 3D input device. Intersecting this ray with the menu plane enables menu selection, see Figure 2 c. We will refer to this method as the RAY method. A disadvantage of this method is that it cannot (or hardly) be used for menus positioned very close to the user. It can however, deal with menus far away. Furthermore, the menus can be manipulated while having the hand with the input device at a comfortable position. In the PSS, the user can rest his elbow on the table while manipulating the menu, thereby reducing arm strain.

3.4. Projective

In the projective methods, we simply project the 3D position of the input device onto the menu plane in 3D space. We have implemented two types of projection: line of sight and orthogonal, see Figure 2d and e. For the latter, the 3D position of the input device is simply projected orthogonally onto the menu plane. We will refer to this method as the ORTHO method. In the line of sight projection, a line going through the 3D input device position and the position midway between the users eyes is intersected with the menu plane to determine the 3D cursor position. This technique is similar to the image plane ‘occlusion’ technique by Pierce et al. [PFC*97]. We will refer to this method as the SIGHT method.

The ORTHO method is not suitable for menus located far away from the user. Such menus will have to be scaled up to be readable. Therefore, large cursor movements (and thus large hand motions) are required to manipulate the menu. The required motions can exceed the user’s reach or the working volume of the tracking system. The SIGHT method does not suffer from this limitation. Furthermore, the SIGHT method allows for lesser arm motion to achieve menu selection; the closer the input device is held towards the eye, the larger the resultant cursor motions will be.

4. Evaluation

4.1. Set Up

We have evaluated the five different menu selection techniques in a desktop virtual reality system, the Personal Space Station [ML02]. The display of the PSS consisted of a 22" IIYama CRT monitor with a screen size of 40 x 30 cm and configured at 1280 x 1024 pixels at 120 Hz. The reflected display of the PSS, called the Virtual Focus Plane (VFP), was perceived by the user at a depth of about 50 cm. The VFP faced the user approximately perpendicular. Stereoscopic viewing was accomplished with the use of a NuVision polarization screen and matching glasses. For tracking the user’s head, we used a custom optical tracking system based on two iBOT FireWire Cameras running at 30 Hz [MJR03].

For interaction, the PSS was equipped with a standard Logitech desktop mouse, a 3D stylus-like input device, and a foot switch. The 3D input device was tracked using a custom optical tracking system based on infrared lighting, retroreflective markers, and two progressive scan LeutronVision cameras running at 60 Hz [LM03].

The user test consisted of a series of menu selection tasks,

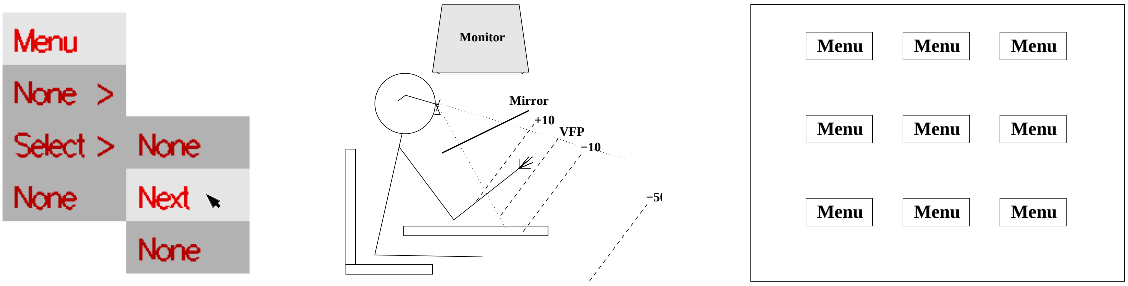


Figure 3: *Left: The menu. Center: The menu plane positions at different depths Right: The menu positions.*

one for each technique. The menu is depicted in Figure 3 a. The position of the menu was varied. Four planes were chosen, with varying distances (depth) from the subject: a plane close to the subject in front of the VFP, the VFP itself, a plane 10 cm behind the VFP, and a plane 50 cm behind the VFP, see figure 3 b. Nine different menu locations on each plane were used, see Figure 3 c. The size of the menu was adjusted for each plane such that it would always cover approximately the same screen area.

For each selection method, three series of nine menu selections had to be performed for each plane. For the POS and ORTHO methods the plane furthest away from the subject was omitted, since menus at these distances cannot be manipulated with these methods. For the RAY method, the plane nearest to the subject was omitted. For each series of nine selections, the menu positions were randomly distributed over the nine locations on the plane. In addition, the plane order and menu method order were varied between subjects. In total, each subject performed 459 menu selections, 81 for each of the POS, ORTHO, and RAY methods (3 planes with 27 positions each), and 108 for each of the MOUSE and SIGHT methods (4 planes with 27 positions each).

Six subjects participated in the experiment. All subjects were male, right handed, and with normal or corrected to normal vision. All subjects were familiar with the PSS. Before a session, the different methods were explained to the subject, and the subject performed a number of menu selections to practice each method and make sure he understood the concept. After these practice runs, the subjects were asked to perform a series of menu selections for each method. Before a series for a particular method, the subjects were given a number of trial selections to get acquainted with the particular method again. During the sessions, subjects were encouraged to speak out loud while they performed the tasks, but were instructed not to pause during a method session. In between method sessions, they were allowed to pause until they indicated they were ready for the next session. After the complete session, the subjects were asked to sort the methods according to their preference, and provide their motivation.

4.2. Results

Figure 4 depicts a box plot of the menu selection times for each method over all subjects, along with a summary of the selection times per method. An analysis of variance (ANOVA) showed that menu technique was indeed a significant factor ($p < 0.01$). No significant learning effects were found within the methods or subjects.

Figure 5 depicts a box plot of the menu selection times for each method over all subjects per depth plane, and a summary of the results. An analysis of variance (ANOVA) showed that the plane location was not a significant factor within any of the menu methods.

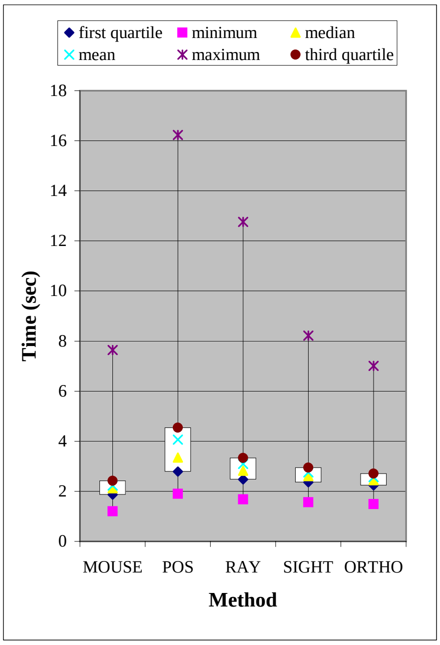
5. Discussion

The MOUSE method proved to be fastest. The POS method proved to be the slowest. The differences in the other methods are less pronounced, but nonetheless significant. ORTHO was the fastest of the three, followed by respectively the SIGHT and the RAY methods.

We are not very surprised that MOUSE was the fastest, considering that all subjects were very experienced with using the standard desktop mouse and the difference between using the mouse for menu selection in the MOUSE method and in a standard WIMP interface is marginal.

We also expected the POS method to be the slowest, since it requires precise movements in three dimensions instead of two dimensions for the other methods. Furthermore, it requires the user to have an accurate depth perception of both the menu and the cursor to position the cursor on the menu. We noticed most subjects had to actively search for the correct depth by moving the cursor back and forth until they hit the menu. This was particularly noticeable when the menu switched between depth planes.

We were surprised by the order of the other methods though. The ORTHO method proved to be faster than the SIGHT method, despite the extra amount of movement that was required in the ORTHO method. Possible causes include the fact that the SIGHT method is more sensitive to noise in the interaction and head trackers. In addition, it requires a



	Mean	Standard Deviation
MOUSE	2.23	0.59
POS	4.06	2.11
RAY	3.09	1.10
SIGHT	2.75	0.69
ORTHO	2.56	0.60

Figure 4: *Menu selection times per method.*

more controlled hand movement because of the larger cursor displacement. The same applies to the RAY method; tracker noise and an unsteady hand have large effects on the selection.

We were also surprised by the fact that for none of the methods the menu plane position (in depth) was a significant factor. We had expected that the plane located far away (50 cm behind the focal plane) might have been more difficult because of possible fusion problems (objects located at that plane show a large parallax when rendered on the display screen). In fact, one of the subjects specifically complained about that. In this regard, we also expected menus rendered on the focus plane (zero parallax) to have a positive effect on selection time. Apparently, even when fusion could be difficult, menu selection times do not necessarily suffer.

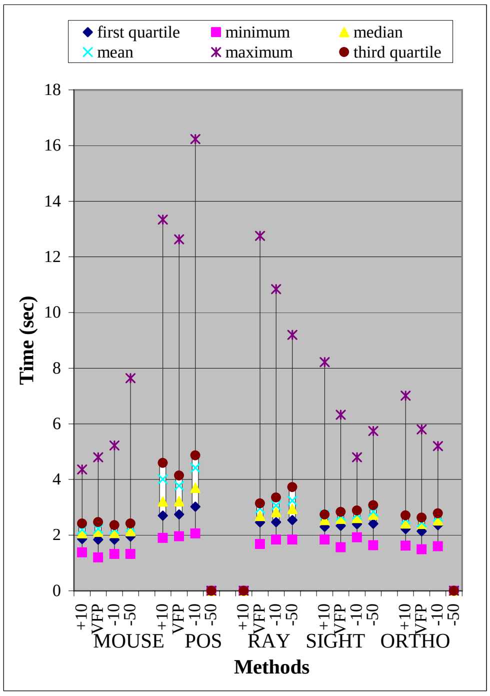
The subjects were asked what their favorite method was. They all agreed on the MOUSE method. They also agreed on the POS method being the least favorite. There was no consistent ordering of the other methods among the subjects. One subject commented that out of the other three methods he liked the ray method best because he could rest his elbow on the table while manipulating the menus.

All subjects complained about arm strain for the methods that used the 3D input device. A session for a single menu method with that device could easily take over 5 minutes, during which the subjects mostly kept their arm in the air. Except for the one subject mentioned earlier, none of the subjects took the opportunity to rest their elbow on the table for the RAY method. It was observed that instead of using wrist movement to point the ray towards the menu, most subjects kept the ray at a steady angle and moved their entire fore-arm to perform the menu selection. Only one subject complained about arm strain during the use of the MOUSE method.

6. Conclusion

In this paper, we have evaluated five different menu selection methods for desktop or dexterous virtual environments. Results have shown, that using the standard desktop mouse provides the fastest menu selection times. Subjects also preferred using a standard desktop mouse over using a 3D input device. Using the exact 3D position of a 3D input device for menu selection performs worst, and user preferences also indicate this method to be the least favorite. The difference in performance of the other three selection methods is less pronounced.

We conclude that it is best to use the standard 2D desktop mouse for menu selection in 3D desktop systems. However, when working with a 3D device and a device swap is undesired, menu manipulation can be performed with one of the projective methods or the ray casting methods. If many menu selections have to be performed, the ray selection method should probably be used since it allows the user to keep his arm rested on the table and therefore reduces arm strain. If menu selection is sporadic, it is best to use the orthogonal projection method, unless the location of the menu is far away from the user. In that case, the line of sight projection method is to be preferred.



		Mean	Standard Deviation
MOUSE	+10	2.19	0.54
	VFP	2.22	0.55
	-10	2.21	0.58
	-50	2.29	0.67
POS	+10	4.00	2.16
	VFP	3.77	1.76
	-10	4.41	2.33
RAY	VFP	2.96	1.09
	-10	3.07	1.06
	-50	3.24	1.12
SIGHT	+10	2.73	0.89
	VFP	2.69	0.61
	-10	2.73	0.58
	-50	2.84	0.63
ORTHO	+10	2.56	0.66
	VFP	2.45	0.54
	-10	2.66	0.58

Figure 5: *Menu selection times per plane per method.*

As for future work, we plan to develop new interaction techniques and metaphors specifically aimed at 3D desktop or dexterous VR and AR environments. We plan to compare such interaction techniques with other techniques developed for HMD or projection based, fully immersive VR and AR environments, as well as with the menu techniques described in this paper.

References

[ABW93] Arthur K., Booth K., Ware C.: Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems 11*, 3 (July 1993), 239–265. 2

[AW00] Arsenault R., Ware C.: Eye-hand co-ordination with force feedback. In *CHI 2000 Conference Proceedings* (2000), pp. 408–414. 2

[BH95] Bowman D., Hodges L.: *User Interface Constraints for Immersive Virtual Environment Applications*. Technical Report GIT-GVU-95-26, Georgia Institute for Technology, 1995. 2

[BW01] Bowman D., Wingrave C.: Design and evaluation of menu systems for immersive virtual environments. In *Proceedings of the IEEE Virtual Reality Conference 2001* (2001), pp. 149–156. 2

[Dee92] Deering M.: High resolution virtual reality. In *Computer Graphics (SIGGRAPH '92 Proceedings)* (1992), Catmull E., (Ed.), vol. 26, pp. 195–202. 1

[JE92] Jacoby R., Ellis S.: Using virtual menus in a virtual environment. In *Proceedings of SPIE: Visual Data Interpretation* (1992), pp. 39–48. 2

[KL04] Kok A., van Liere R.: Co-location and tactile feedback for 2D widget manipulation. In *Proceedings of the IEEE Virtual Reality Conference 2004* (2004), pp. 233–234. 2

[LM03] van Liere R., Mulder J.: Optical tracking using projective invariant marker pattern properties. In *Proceedings of the IEEE Virtual Reality Conference 2003* (2003), pp. 191–198. 3

[LSH99] Lindeman R., Sibert J., Hahn J.: Hand-held windows: Towards effective 2D interaction in immersive virtual environments. In *Proceedings of the IEEE Virtual Reality Conference 1999* (1999), pp. 205–212. 2

[MBS97] Mine M., Brooks F., Sequin C.: Moving objects in space: Exploiting proprioception in virtual-environment interaction. In *Computer Graphics (SIGGRAPH '97 Proceedings)* (1997), vol. 31, pp. 19–26. 2

[MJR03] Mulder J., Jansen A., van Rhijn A.: An affordable optical head tracking system for desktop VR/AR systems. In *IPT/EGVE 2003 Proceedings* (2003), pp. 215–223. 3

[ML02] Mulder J., van Liere R.: The personal space station: Bringing interaction within reach. In *VRIC 2002 Conference Proceedings* (2002), pp. 73–81. 1, 3

[PFC*97] PIERCE J., FORSBERG A., CONWAY M., HONG S., ZELEZNIK R., MINE M.: Image plane interaction techniques in 3D immersive environments. In *1997 Symposium on Interactive 3D Graphics* (1997), Spencer S., (Ed.), pp. 39–43. 2, 3

[PS96] POSTON T., SERRA L.: Dextrous virtual work. *CACM 39*, 5 (1996), 37–45. 1

[SES99] SCHMALSTIEG D., ENCARNAÇÃO L., SZALAVÁRI Z.: Using transparent props for interaction with the virtual table. In *Symposium on Interactive 3D Graphics* (1999), pp. 147–153. 2

[SPHC95] SERRA L., POSTON T., HERN N., CHOON C.: Interaction techniques for a virtual workspace. In *Proceedings ICAT/VRST '95* (1995), pp. 79–80. 2

[WAB93] WARE C., ARTHUR K., BOOTH K.: Fish tank virtual reality. In *INTERCHI '93 Conference Proceedings* (1993), Ashlund S., Mullet K., Henderson A., Hollnagel E.,, White T., (Eds.), pp. 37–42. 2

[WF96] WARE C., FRANCK G.: Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics 15*, 2 (April 1996), 121–140. 1

[WR99] WARE C., ROSE J.: Rotating virtual objects with real handles. *ACM Transactions on CHI 6*, 2 (1999), 162–180. 2