Int. J. Human–Computer Studies (2000) **52,** 111–129
Article No. ijhc.1999.0326
Available online at http://www.idealibrary.com on IDEAL®

Navigation strategies with ecological displays

CATHERINE M. BURNS†
Cognitive Engineering Laboratory, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada

(*Received 22 December 1998 and accepted in revised form 7 June 1999*)

Ecological interface design (EID) has shown success as an approach for interface design in the case of a process control microworld. However, in applying the EID approach to larger systems, questions arise as how to support the navigation and integration of abstract information. In this study, three ecological displays were developed for a simulated power plant from the same abstraction hierarchy. The displays differed in the integration of abstract information, demonstrating high-space low-time, low-space high-time, and high-space high-time integration. While using the displays, the screen actions of subjects were recorded and their navigation movements studied through maps of navigation trajectories. Distinct differences were apparent between the temporally integrated and the temporally separated displays. In the temporally separated displays, clear scanning patterns emerged and these scanning patterns were correlated with improved performance on the display. This suggests that scanning patterns are an adaptation to needed but separated information. It also suggests that functional integration is an important characteristic to support when designing large ecological displays.

1. Introduction

Increasingly, computer displays are being applied to large and complex industrial systems. Due to the scale of these systems, a new class of human factors problem is arising. These new questions involve how to support the operator in controlling these large systems, how to maintain integration across many displays and how to support effective navigation.

Ecological interface design (EID) is one approach that has been shown to have considerable success in supporting problem-solving activities in process control situations. EID begins with a needs analysis of the work domain, the object or system for which the interface is being designed. This analysis looks for physical capability information, designed for purpose information, and functional information linking the physical implementation to the designed for goals. In general, this results in an interface with several levels of information beyond equipment settings and values. This additional information helps the user to solve the problem, taking advantage of a reasoning path from goal to physical equipment.

†Address for correspondence. Catherine M. Burns, Advanced Interface Design Laboratory, Department of Systems Design Engineering, University of Waterloo, Waterloo, Ont. Canada N2L 3G1. email: c4burns@uwaterloo.ca.

1071-5819/00/010111 + 19 $35.00/0

The most thoroughly studied implementation of EID uses the DURESS II simulation (Vicente, 1992). The DURESS II simulation is a process control simulation with two pumps, a network of pipes and six valves, ending in two heated tanks. The traditional purely physical display, representative of current engineering interfaces, contains only information on the components and system outputs. The ecological display, in contrast, gives the operator information on mass flows and mass and energy relations. The ecological display for DURESS II has shown consistent success in improving problem solving and effective control in several studies (see e.g. Vicente, 1992; Vicente, Christoffersen & Pereklita, 1995; Pawlak & Vicente, 1996). The DURESS II simulation, however, does not represent the scale of an industrial plant. It is important, therefore, to continue to explore the ability of EID to support problem solving in larger systems.

EID advocates that operators of complex plants require an interface which gives them a complete picture of how the plant meets its designed for purpose. It begins with an analysis of the plant, in terms of its purpose, and how it meets that purpose at different levels of abstraction. This analysis results in the creation of an abstraction hierarchy (Rasmussen, 1985). In comparison with traditional displays, EID provides additional information to the operator, primarily at higher abstraction, or "functional" levels. When applied to larger systems, therefore, it could be anticipated that EIDs would be prone to problems of integration. The integration of large EIDs could be more challenging due to additional information presented. Furthermore, any loss of integration between the functional levels could detract from the benefits of EID. It is important, therefore, that EID be studied with larger systems, and in situations where all the information cannot be presented in one view. Studies under these conditions are needed before EID can be used as an industrial design approach for computer interfaces.

As a basis for understanding how large EIDs should be integrated, we examined what is known about how people use EIDs to reason through the abstraction hierarchy. Two of the DURESS studies have looked specifically at how people reason within the abstraction hierarchy while problem solving. Vicente *et al.* (1995) conducted a verbal protocol study of fault diagnosis with an early non-interactive version of the DURESS simulation and found that more frequent verbal indications of functional information correlated with improved fault diagnosis performance. Janzen and Vicente (1998) separated the abstract information of the DURESS display into four separate views, one for each level of the abstraction hierarchy. Similarly, they found that more frequent visits to the abstract function level corresponded with better fault diagnosis performance on normal and fault trials. Both these studies show the importance of maintaining accessibility to the higher functional levels, particularly abstract function, to support fault diagnosis behaviour.

Based on these ideas, we hypothesized that the most critical dimension of integration of EIDs would be along the dimension of abstraction. It would be critical to create large EIDs where users would continue to access higher functional levels in order to continue to gain the problem-solving benefits of EID. For this reason, we obtained a medium-scale power plant simulation with over 400 variables. We created an abstraction hierarchy model of this simulated plant and then developed views consistent with that abstraction hierarchy. Then these views were integrated along the abstraction dimension in one of three different ways, high-spatial low-temporal integration, low-spatial high-temporal integration and high-spatial high-temporal integration.

We hypothesized that the highest degree of spatial and temporal integration would provide the best performance on a fault detection and diagnosis task. We tempered this hypothesis with a counter-hypothesis that the benefits of this approach would degrade as the display became more cluttered. In terms of using the displays, we anticipated that the three display types would be used differently, and result in different navigation approaches. This paper concentrates not on the fault management performance with these displays but on the differences in navigation strategies that they generated.

In general, navigation in computer interfaces has been sparsely studied and rarely in the context of an industrial system of this scale. The result of this work, therefore, should be of interest to computer interface designers in general and give insight beyond the field of EID.

2. Method

This section discusses the simulated power plant, the work domain, used for this experiment and the basics of the abstraction hierarchy model that was used to design the displays. The subject pool and experimental task are outlined.

2.1. THE WORK DOMAIN

A simulated ABB conventional power plant was chosen as the domain for the displays. The simulation, consisting of over 400 variables, was considered to be of adequate scale to create displays complex enough to invite navigation actions to occur. A controlled evaluation of navigation in an EID for a work domain of this scale, to our knowledge, has not been previously reported in the literature. Furthermore, this simulation was an industrial research level simulation with realistic plant dynamics and scenarios, lending external validity to this work.

2.2. APPARATUS

The on-line portions of the experiment were conducted on a Dell Pentium 133 MHz computer with 32 Mb RAM and a 15 in SVGA monitor. The simulation of the power plant was provided by ABB and was proprietary ABB software. The displays for the simulation were created using Fix 6.0, a commercial supervisory control and data acquistion program with an interface development module. All screen actions were recorded on videotape using an Extron Super Emotia scan converter connected to a Sony video tape recorder (VTR).

2.3. ABSTRACTION HIERARCHY-BASED DISPLAYS

An abstraction hierarchy model was made of the plant. The abstraction hierarchy model had 10 cells covering four abstraction levels and four aggregation levels, as shown in Figure 1. All 400 plant variables were involved in the abstraction hierarchy. Because these models are quite extensive, they have been reported elsewhere (Burns, 1998). Briefly, four levels of abstraction were modelled. At the level of functional purpose (FP), the models contained information on plant goals and outputs. At the level of abstract

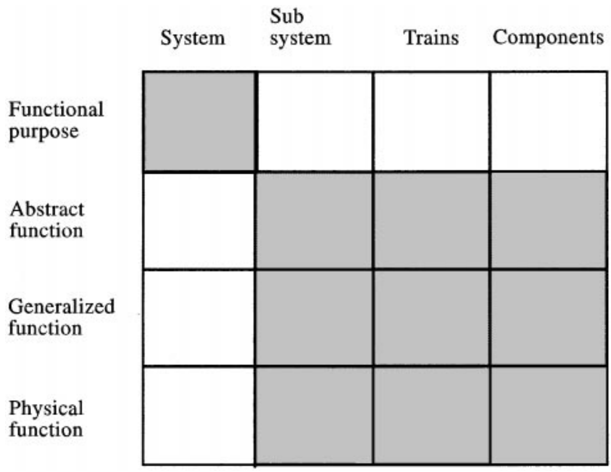

FIGURE 1. Modelled cells of the abstraction hierarchy.

function (AF), mass and energy flows throughout the plant were represented. The generalized function (GF) level contained models of plant processes, temperature, pressure and state changes. The physical function (PF) level contained models of plant equipment, its settings and behaviour. Four levels of decomposition were used. The system level represented the plant as a unit, the sub-system level showed the major units of the plant (e.g. turbines). “Trains” referred to redundant parallel lines of components and provided an intermediate level of decomposition between sub-systems and components. Since the plant was simulated, no cells were modelled at the level of physical form since this information was not available from the simulation. Physical form information typically is information on the condition and appearance and location of plant equipment, which is not available in a software simulation of a plant.

Based on the abstraction hierarchy model of the plant, 10 views were designed to demonstrate the information in each cell of the abstraction hierarchy. These views were then integrated in three different ways (Table 1). All integrations were performed along the dimension of abstraction and no integration was performed along the decomposition dimension. One display demonstrated temporal, but not spatial, integration, *a low-space high-time* (LH) display (Figure 2). This display showed all four abstraction levels at the same time, but separated into four windows. For example, boiler energy, boiler temperature and boiler level, three pieces of information at different levels of abstraction, would be visible simultaneously but in different windows. A second display implemented spatial integration, but not temporal integration, a *high-space low-time* (HL) display (Figure 3). This display showed one abstraction level at a time. Following the boiler example, only boiler level would be visible until the user changed the display. The third display showed minimal separation, or the tightest possible spatial and temporal integration of the abstracted information. This display was a *high-space high-time* (HH) display (Figure 4). In this display, information at all levels of abstraction was visible at the same time, and closely integrated in spatial location. Boiler level, boiler temperature and boiler energy information would all be brought together in close proximity.

TABLE 1
Display conditions

Condition	AH levels shown	Number of windows
High-space low-time	1	1
Low-space high-time	4	4
High-space high-time	4	1

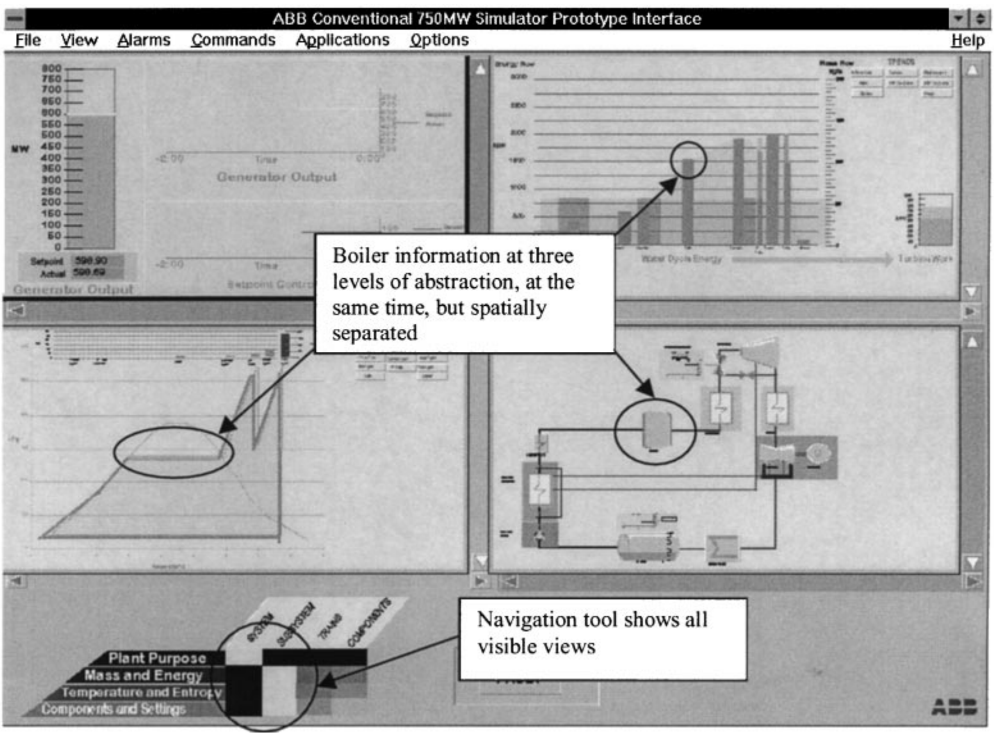

FIGURE 2. Low-space high-time display.

It could be argued that a "low-low" display was required. A "low-space low-time display" might display the information in four separated windows, like the LH display, but then only make one of those windows visible at a time. When considering this display as an experimental option, we decided that this display made such poor use of screen real estate that, although possible, it was not a feasible design alternative and in conflict with our goals of having a valid and industry-relevant study.

All three displays showed a navigation tool (Figure 5) that displayed which cell(s) of the abstraction hierarchy was visible. The navigation tool was located in the lower left-hand corner of the screen. The abstraction hierarchy structure was used as a navigation map to show the current location in the hierarchy. This tool could also be used to

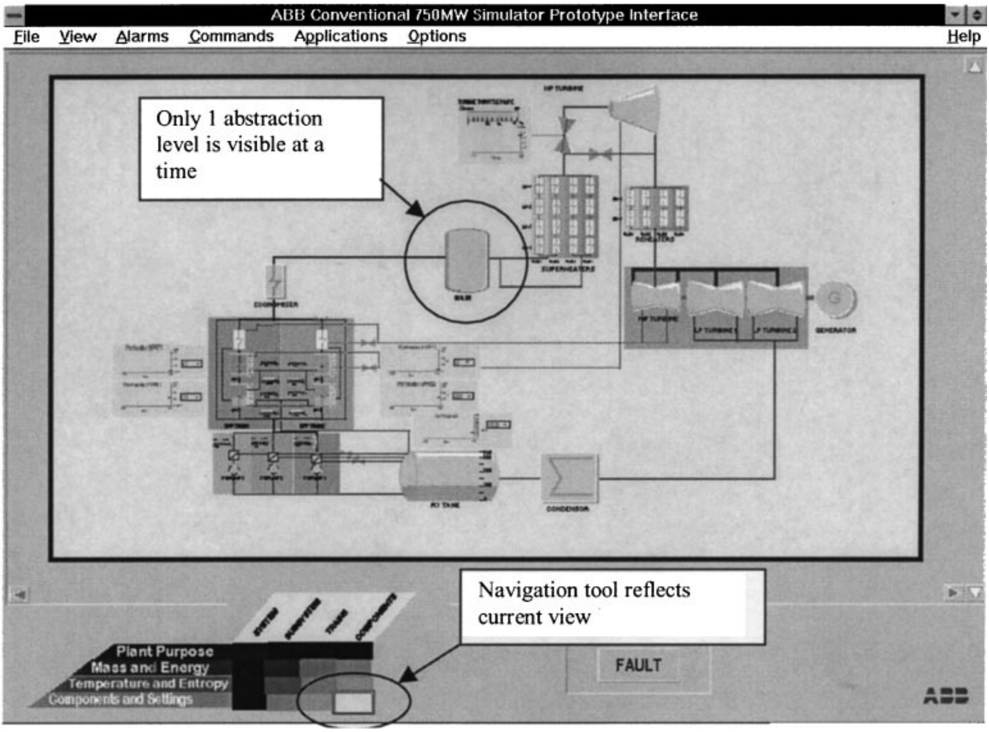

FIGURE 3. High-space low-time display.

make changes in the display views by clicking on new abstraction hierarchy cell. In all conditions, subjects were able to display any view of the abstraction hierarchy at any time. The navigation tool was identical for all displays, except that in the temporally integrated displays it showed, correctly, that multiple views were visible. Figure 5 shows the tool for the temporally integrated views.

2.4. SUBJECTS

Eighteen subjects were recruited from the third year class of Mechanical and Industrial Engineering at the University of Toronto. The subjects ranged in age from 19 to 24 years with the mean being 20.9 years ($s = 1.2$). Six subjects were female and 12 were male. All subjects had taken exactly one course in thermodynamics ($s = 0$) and, in fact, had been in the same class. Subjects were tested on the Spy Ring History Test and subject groups were balanced based on their holism scores. This was because tendencies towards holism have been found to interact positively with performance on ecological displays (Torenvliet, Jamieson & Vicente, 1998). After the criterion of balancing the groups was applied, the subjects were randomly assigned to display groups. Each group had six subjects and this was a between-subjects design. To maintain anonymity, each subject was given the

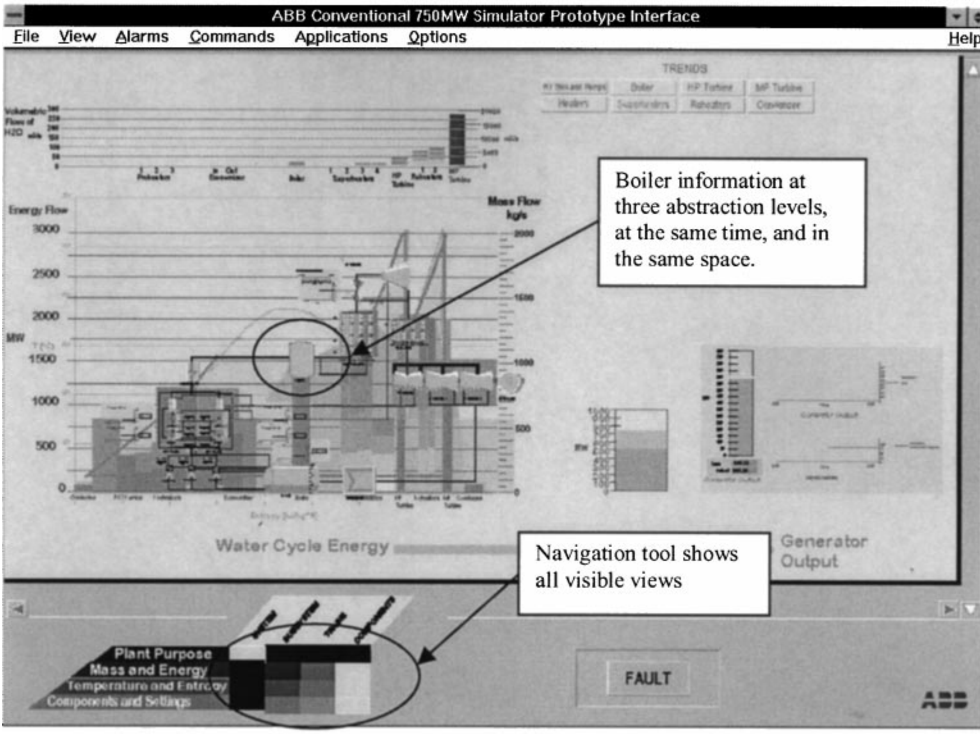

FIGURE 4. High-space high-time display.

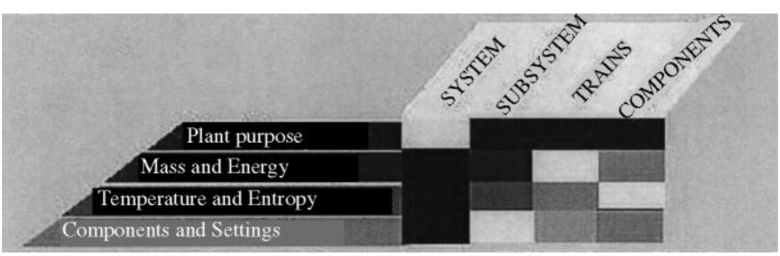

FIGURE 5. The navigation tool for temporally integrated views.

name of a famous explorer. It is by these names that subjects have been identified in this paper.

Subjects received 6 h of training on the plant and the displays and were expected to demonstrate a high level of understanding of the plant and competence with the displays before proceeding in the experiment. To confirm competence, subjects were tested on the plant and the displays. They were expected to answer 80% of the questions correctly in order to proceed with the experiment.

2.5. EXPERIMENTAL TASK

The primary task of the experiment consisted of monitoring the simulation during 17 trials. During these trials, the plant would operate normally for a short, randomly determined period time. Following this period of normal conditions, either an equipment failure would occur, or the plant would continue to operate normally. Equipment failures and normal scenarios were interspersed randomly, so subjects would not anticipate an equipment failure and, therefore, respond prematurely. Equipment failures and normal scenarios were randomized once and then presented to all subjects in the same order of presentation. Most scenarios had a maximum time of 15 min beyond which point, if no diagnosis had been made, subjects were prompted to offer their best diagnosis and proceed to the next trial. Subjects were asked to monitor the plant and, if they detected a fault, to indicate that detection. They were then expected to proceed to diagnose the fault. The time to detect a fault, the time to diagnose a fault and the fault diagnosis were recorded. All screen actions were captured on videotape using the scan converter.

3. Results

From the analysis of screen actions, navigation strategies were derived from those actions. Fault management performance results are discussed briefly, but have been published elsewhere (Burns, in press). In particular, from the screen actions that were collected, navigation trajectories, or paths, made by the subjects as they worked through the display space were mapped out for each display. These strategies were collected to see whether or not the different integration approaches modified the process by which users manipulated the displays. Marked differences in the frequency of navigation actions were apparent between the displays. To quantify these differences, we examined the dwell time of the subjects in the various cells of the display space and then correlated dwell time with monitoring performance. Detection time was chosen as the measure of monitoring performance.

3.1. FAULT MANAGEMENT PERFORMANCE MEASURES

On the fault management measures of detection time, diagnosis time and diagnosis accuracy it was found that the displays differed in how well they supported these aspects of the fault detection task (Table 2). The HL display generated the fastest detection times, whereas the HH display tended to generate the fastest diagnosis times and the most frequently correct diagnoses (Burns, in press). The general conclusion from these results was that, while the HL integration supported the monitoring task better, the tighter

TABLE 2
Significant performance measure results

Measure	Best performing display
Detection time	HL
Diagnosis time	HH
Diagnosis accuracy	HH

space–time integration of the HH display supported the problem-solving aspects of the task better. Burns (1998) contains a fuller discussion of these results.

It appeared from these results that HH integration was the best alternative for supporting problem solving, but, possibly because of the busy and cluttered nature of the display, it did not support monitoring activities as well. Interestingly, the LH "windowed" situation, a commonly seen design, was not a promising design alternative from the standpoint of any of the measures. Based on these performance differences, we decided to examine how the displays were actually used to see if that would give insight to the performance differences.

3.2. TRAJECTORIES OF NAVIGATION ACTIONS

To evaluate navigation actions made with the three displays, the videos collected were analysed. Movements from cell to cell of the display hierarchy were recorded. In order to compare how well the three displays supported monitoring actions, movements were recorded from the start of the trial unit the subject indicated detection of a fault. The subjective indication of a fault was taken to be the end of monitoring actions because, from the point onward, the subject's actions could be expected to be directed towards diagnosing the fault, and no longer demonstrating general monitoring behaviour. In addition, the *subjective* indication of a fault was used in place of the *actual* start of the fault since it was possible for subjects to perceive (wrongly) that a fault was occurring when no fault had occurred.

Movements between the displays were recorded and then mapped against a background of the possible views. Recall that the hierarchy of views was a direct mapping of the abstraction hierarchy used to model the plant. Each movement was indicated as an arc, spanning clockwise from the first view to the next selected view. Figure 6 shows this

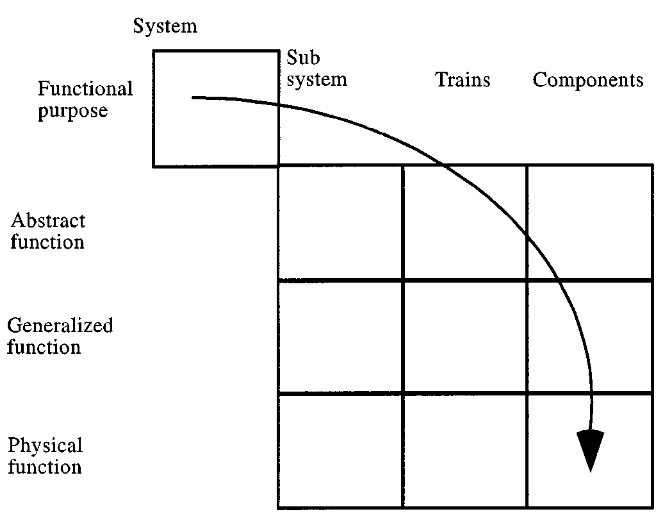

FIGURE 6. Background and movement for trajectory diagrams.

background and an example of a movement from FP to PF components. Movements made more than once were indicated by the thickness in point value of the line.

These trajectories were plotted for two normal trials, that is, trials where no plant failure occurred. These trials were chosen because they provided a longer time period during which monitoring actions could be observed. In addition, one of these trials occurred early in the experiment, at trial 4 of 17. The other trial occurred late in the experiment, at trial 14 of 17. This presented a good opportunity to examine any changes in monitoring strategies as the experiment proceeded and the subjects adapted to the displays. The following figures, Figures 7, 8 and 9, present the trajectories of each subject on trial 4 (left) and trial 14 (right). Figure 7 shows the trajectories for the HL display, Figure 8 for the LH and Figure 9 for the HH display. To allow the trajectories to be compared on the same page, extraneous information has been removed from the plots and Figure 6 is recommended as a template to indicate the cells of the abstraction hierarchy map.

The trajectories of the two temporally integrated displays, LH (Figure 8) and HH (Figure 9), look different from the HL trajectories (Figure 7) for two reasons. First, since every abstraction level was visible at the same time, the subject's movements were constrained to changes in level of decomposition. The arcs, therefore, only span decomposition levels. Second, these subjects started from an envelope of views, not from a single starting view. For this reason, the starting views have been indicated by a dot in the centre of the view.

Despite this, certain differences between the trajectories are clearly apparent

(1) The overall number of navigation actions taken by subjects in the HL display were much greater than in the other two display situations.
(2) Navigation actions taken by the subjects in the HL display tended to span all four abstraction levels.
(3) Navigation actions taken by the subjects in the HL display did not span the decomposition levels.
(4) The trajectories from the temporally integrated displays (LH, HH) appeared to be very similar.
(5) From trial 4 to 14, subjects using the HL display tended to reinforce their scanning pattern. The scanning pattern, regardless of form, tended to become less variable, yet was repeated more frequently.

While differences were clearly apparent from observation of the trajectories, we proceeded to seek quantifiable measures that might reflect some of the differences in the strategies that were observed.

3.3. QUANTIFICATION OF STRATEGIES

In order to quantify these strategies for comparison, measures were made of the average time spent per cell, or *dwell time*. This measure was adopted from Vicente *et al.* (1995) and Janzen and Vicente (1998) who used similar measures to quantify trajectories through an abstraction hierarchy space. Used as an overall measure, dwell time would differentiate between highly active patterns and less active patterns. Examined cell by cell, it would indicate where subjects spent long visits or more fleeting visits. Although the navigation

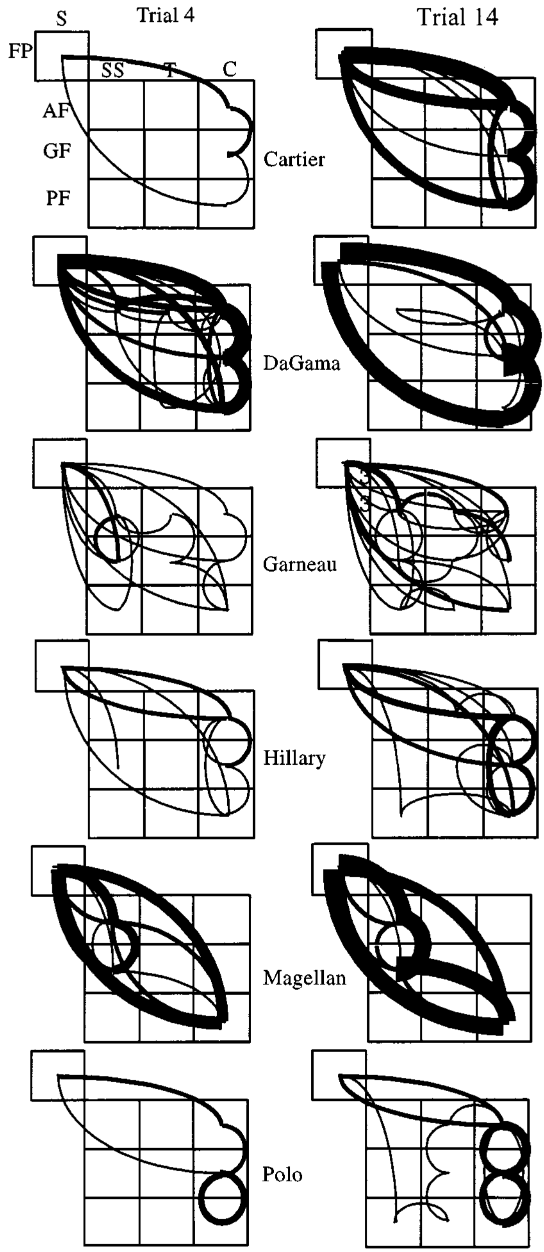

FIGURE 7. Navigation trajectories for HL display.

maps gave a rich picture of number of visits and where subjects went, a time-based measure was required to confirm that the distribution of those visits was similar, or different, across the different display types.

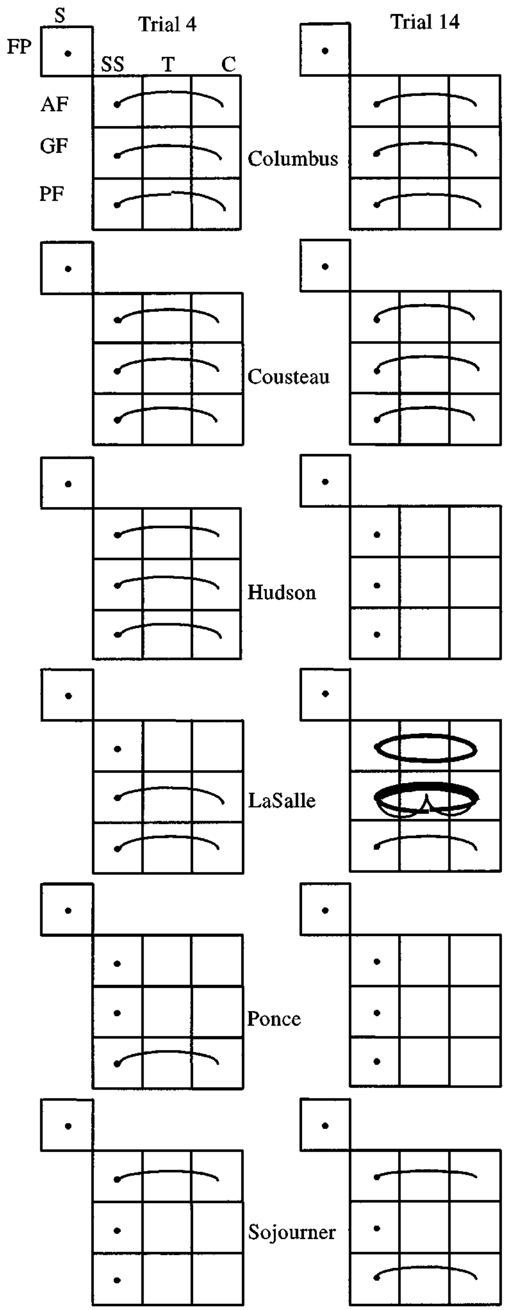

FIGURE 8. Navigation trajectories for LH display.

3.3.1 Overall dwell time

The time spent per cell per visit, or *dwell time* (Janzen & Vicente, 1998) was recorded for all scenarios. Dwell time was calculated overall by display, and then, within displays, by

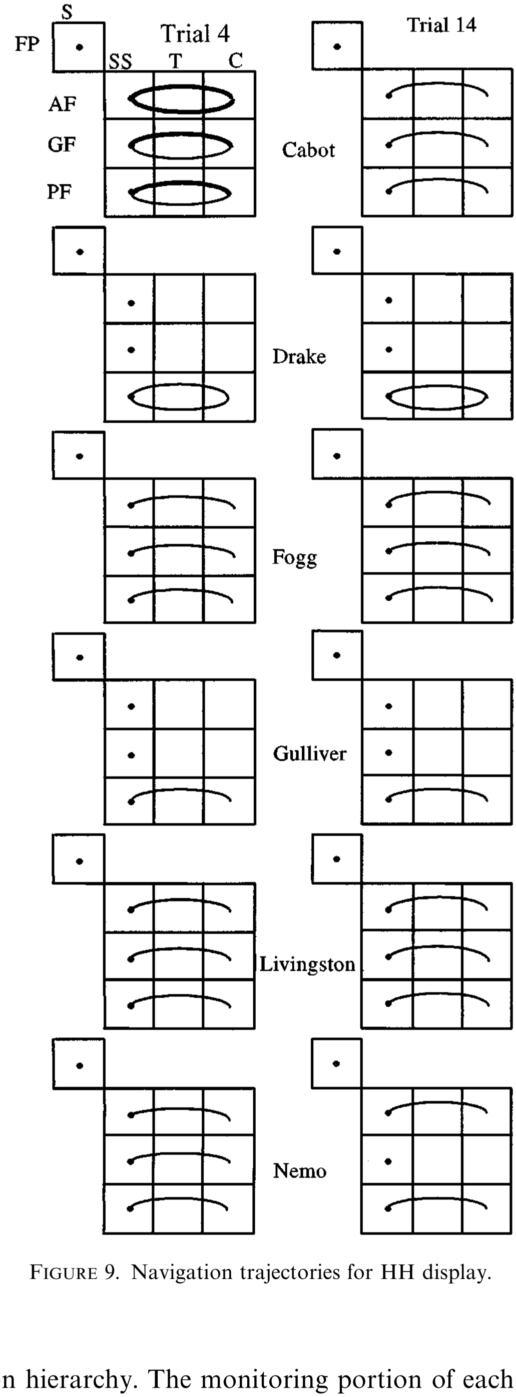

FIGURE 9. Navigation trajectories for HH display.

cell of the abstraction hierarchy. The monitoring portion of each of the scenarios was used in the calculation of this measure. Figure 10 shows the mean overall dwell time by display group, with error bars of one standard deviation.

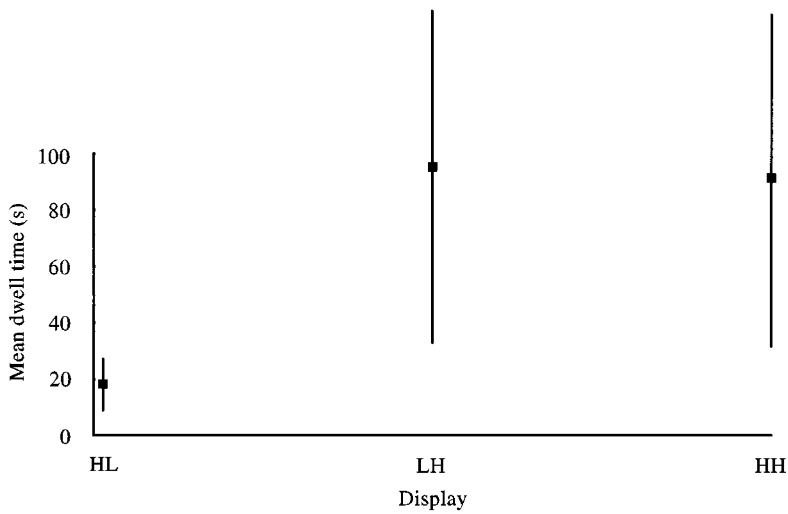

FIGURE 10. Overall dwell time by display type.

These means were analysed by an ANOVA model with Display as a fixed effect and Subject as a random effect nested within Display. The differences between the means were statistically significant [$F(2, 15) = 6.06$, $p = 0.01$]. The size of the effect was quite large. The HL display had a mean dwell time of 17 s compared to 86.9 s and 87.4 s for the LH and HH displays, respectively. In conjunction with the navigation maps, this confirmed that subjects using the HL display showed more rapid movement between display cells. Presumably, this behaviour was a response to the temporal separation. However, it could also be that, having less information available per view, there was less need for the subjects in this display group to spend a lot of time to understand the display.

3.3.2. Dwell time per view

Dwell time per view was calculated only for the AF, GF and PF views. The dwell time on the FP view was not compared since this view was always visible on the temporally integrated displays and hence, the comparison would be unfair. In addition, since in the LH and HH displays, all abstraction levels were visible, these data only show the distribution of dwell time across the decomposition levels at each abstraction level. The data for the HL display were calculated to show the same distribution so that comparisons could be made between three displays. These data were analysed using the same ANOVA model used for the mean dwell time. Table 3 presents the means for each display and the significance level resulting from the statistical analysis. All times are in seconds. A "†" indicates that no mean could be calculated for that view since no subjects visited that view. An asterisk has been used to indicate statistically significant results at the $p < 0.05$ level. There are occasionally missing observations in the data set as not all subjects used all views.

TABLE 3
Mean dwell time per cell by display

View	HL	LH	HH	F-test
AF/sub-system	9.51	56.63	57.05	$F(2, 14) = 0.61, p = 0.51$
AF/trains	9.91	29.75	5.00	$F(2, 4) = 1.50, p = 0.32$
AF/component	22.21	125.89	126.52	$F(2, 13) = 37.50, p < 0.01$*
GF/sub-system	8.32	72.95	58.08	$F(2, 13) = 0.80, p = 0.43$
GF/trains	11.00	57.00	6.00	$F(2, 5) = 1.90, p = 0.23$
GF/component	15.17	102.25	121.96	$F(2, 13) = 34.30, p < 0.01$*
PF/sub-system	6.70	40.56	64.68	$F(2, 13) = 0.79, p = 0.47$
PF/trains	12.13	†	†	†
AF/component	20.28	142.42	144.82	$F(2, 14) = 11.10, p < 0.01$*

*Significant result $p < 0.05$.
† Insufficient data.

Several tendencies were confirmed from this analysis.

(1) In the temporally integrated, LH and HH displays, subjects were spending most of their time at the component level views of the displays.
(2) The patterns of use of the views were nearly identical for the LH and HH displays, with almost the same amount of time spent on the same views.
(3) The PF/trains level view was poorly visited and perhaps not a useful view.

Since the navigation maps show only trajectories and not time spent at the views, these results added to the results suggested by the navigation maps.

3.4. CORRELATIONS WITH PERFORMANCE

To get a sense of how these strategies related to performance, these measures were correlated with detection time. Detection time was chosen since it is the performance measure most immediately related to monitoring and these strategies only cover monitoring strategies. Once subjects had detected a fault, we assumed that their activities would be directed towards searching for that specific fault and, from that point onward, would be dependent on the particular scenario. The correlation was examined for overall dwell time and for dwell time on the individual views by display type. Table 4 shows the correlations by display. The probability of the correlation is shown in parentheses following the correlation coefficient. Strong correlations ($r > 0.4$) with low probability ($p < 0.1$) are indicated by an asterisk. For all correlations, $N = 6$ and $\mathrm{df} = 4$. It should be remembered that with these correlations, a positively correlated variable was associated with increased detection time, or poorer performance. A negatively correlated variable was associated with decreased detection time, or better performance.

The first pattern of correlations we observed was that detection time was positively associated with dwell time for the HL display. This meant that, with this display, subjects who spent less time on individual views and made more view changes per minute generate the fastest detection times. This relation, found on the HL display, should be

TABLE 4
Correlations with detection time by display

Measure	HL	LH	HH
Mean dwell	0.8 (0.06)*	– 0.23 (0.66)	0.15 (0.77)
AF/sub-system	– 0.8 (0.05)*	– 0.049 (0.32)	0.26 (0.62)
AF/trains	– 0.66 (0.15)	0.94 (0.01)*	– 0.49 (0.32)
AF/component	0.8 (0.05)*	0.47 (0.35)	– 0.24 (0.65)
GF/sub-system	– 0.75 (0.08)*	– 0.53 (0.28)	0.21 (0.69)
GF/trains	– 0.48 (0.34)	– 0.42 (0.41)	0.17 (0.75)
GF/component	0.77 (0.08)*	0.53 (0.28)	– 0.21 (0.68)
PF/sub-system	– 0.57 (0.23)	– 0.37 (0.48)	– 0.74 (0.09)*
PF/trains	– 0.05 (0.92)	†	†
PF/components	0.47 (0.35)	0.37 (0.48)	0.85 (0.03)*

*$r > 0.4$, $p < 0.1$.
† Insufficient data.

considered in conjunction with the fact that this factor was not correlated with improved detection time on either of the other two displays. This suggests that subjects using this particular display adopted a different strategy to be successful with this particular display. The difference between this display and the other two displays was the degree of temporal integration, suggesting that these subjects may have compensated for the lack of temporal integration by rapidly changing views. Furthermore, the correlation with detection time suggests that this was a successful strategy.

There is another pattern in the strategy correlations that was weak but interesting. It is that increased time spent at the higher aggregation levels may have resulted in improved detection times in both spatially integrated displays. In the HL display, there was a strong negative correlation with AF/sub-system, and again with GF/sub-system with strong positive correlates with the components views. The same pattern occurred with the HH display at the PF level. With the LH display, the pattern is less strong, though the sign of the correlation coefficients is the same. This pattern suggests that the aggregate views may have been better for monitoring and fault detection, but one would need further research to confirm this suspicion.

In summary, these three displays contained the same information but varied significantly in the style of integration that was used. Subjects using the temporally separated display clearly developed a scanning pattern that was distinctive from the other two displays in its travel and frequency. This frequency difference was confirmed through the examination of dwell time in general, and within the different cells of the display space. Furthermore, use of this strategy seemed to correlate with improved performance when using this display.

4. Discussion

This research provides insight into the adaptation of strategies for using different displays. Because all three displays contained the same basic information, the differences seen here can be attributed to the different approaches to integrating the displays. Interestingly, users adapted differently to these different displays.

The obvious counter-argument to this conclusion is that the LH and HH displays simply did not provide the same opportunities for navigation that the HL display provided. This argument is invalid in two ways. First, both LH and HH displays permitted navigation along the dimension of decomposition, however, subjects did not exhibit navigation actions in this direction, even though the opportunity was available. Opportunity for navigation by itself does not invite navigation beyond that of cursory exploration. The second reason this argument is invalid is that, because of the interrelations of the abstraction hierarchy, no navigation actions were required at all in order to detect an equipment failure. An equipment failure would be visible in some form on any of the abstraction views. For example, a break in the boiler would impact boiler temperature and pressure, boiler energy and ultimately plant output. By this argument, navigation actions with even the HL display were not required in order to perform this task. Subjects using the HL display could have solved this problem using the same strategy as subjects using the LH or HH displays. This indicates that the two strategies that emerged in this work are interesting and worthy of some careful discussion.

Two strategies emerged, one of which involved the development of a scanning pattern. This was of particular interest, since other authors, in natural work environments, have noted scanning patterns. Finally, insights were gained on how, and why, scanning pattern develop and the role they play in effective interface use.

4.1. ADAPTATION TO DIFFERENT DISPLAYS

One of the most interesting aspects of this research is that it shows clearly how people adapt to different display environments. Different displays do more than support user performance, but they have an impact on how those users approach their work. Given the three displays seen here, two distinct strategies appeared for using those displays, one a strategy of "sit and watch" on the temporally integrated displays, and the second strategy, one of "scanning" on the temporally separated displays.

4.2. SCAN PATTERNS IN NATURAL ENVIRONMENTS

Scanning patterns have been found in other work. Moray (1986) summarizes the Fitt's group study of pilots' eye movements where scanning patterns have emerged, in this case, with spatially separated information. More recently, Sarter and Woods (1997) have also noted scanning patterns evinced by pilots in field study conditions. Both of these are examples of *spatial scanning patterns* in response to the *spatial separation* of information. Vicente, Mumaw and Roth (1997) noted that power plant operators given many displays with few CRTs tended to scan and flip between displays. This flipping alternates the visibility of displays in time, and therefore demonstrates a *temporal scanning pattern*. In each case, the users are directly compensating for the type of separation with which they have been presented.

4.3. TEMPORAL SCANNING AND THE KEYHOLE EFFECT

Woods (1984) argued that many computerized display systems were plagued by a phenomenon he termed *the keyhole effect*. Presenting the user with a small view (e.g. one

CRT) of a large virtual display space would create the effect of "looking through a keyhole". The keyhole effect is caused by temporally separating display information. In the case of the HL display in this study, users were presented with a keyhole the size of one abstraction hierarchy cell when the entire display space contained the complete abstraction hierarchy of information. By using a scanning strategy, these users compensated for the keyhole effect caused by a small view of a large virtual space (Woods, 1984).

4.4. WHY SCANNING PATTERNS DEVELOP

Interestingly, this work showed how scan patterns develop. Even at the early trial (trial 4) the subject's scan pattern was noticeable. Between trials 4 and 14 the scan patterns tended to maintain their general characteristics, and furthermore, become reinforced. They became stronger in two ways, there was less variation from the primary pattern, and the primary pattern was repeated with increasing frequency.

Scanning patterns, however, do not develop just because displays are separated. In this work, the decomposition views were temporally separated and no scanning developed in this direction. The evidence at this point suggests that people scan in order to view pieces of information that are directly needed to perform their task. The scan patterns seen here provide indirect support for the ecological design framework. If only one level of abstracted information was needed to perform this monitoring task well, surely the users would not have bothered scanning through all four levels.

Furthermore, this scanning pattern, in the case of the HL display, was adaptive. More frequent scanning was associated with improved fault detection performance. The adaptive nature of scanning patterns has been argued by Sarter and Wood (1997), who argued that scanning patterns, in expert pilots, are "highly practiced and fairly effortless" (p. 564). Unfortunately, Sarter and Woods (1997) also argue that situations without scanning patterns require a reflective process requiring mental effort and are, therefore, less desirable. In this study, although scanning improved performance on the HL display, ultimately the best problem-solving behaviour was seen with the HH display. With this display, the subjects did not demonstrate scanning due to the integrated nature of the display. The behaviour-based inference of Sarter and Woods (1997) overlooks two things. First, that scan patterns by themselves are, of necessity, just an adaptation to the separation of information in the environment. Second, they have ignored a third design alternative. It may be possible to design integrated displays that do not require extensive mental effort. This is the direction that research on integrated displays must continue to pursue.

One perspective on these scanning patterns is that, through these scanning patterns, the users were compensating for a flaw in the display. By scanning abstraction levels rapidly, and frequently, the users were compensating for the integration work that was not done by the designer of the display. When information is not temporally integrated, the user must either remember it, or go back and check it. The easier approach of these alternatives was the latter, and the subjects who made the most use of this approach performed the best. Indeed, if these subjects could have scanned the views infinitely quickly, the result would have been the HH display. From the conclusions drawn here, it is anticipated that the spatially separated display (LH) may invite a spatial scanning pattern to emerge. This next dimension indicates a direction for further work.

By examining the process of how displays are used, one can gain insights that extend beyond product measures. This work is novel in that it examined navigation within three similar ecological displays in a controlled experiment, work that has not been done before. This work has demonstrated that, if not carefully integrated, a separated ecological display will invite navigation actions to take place. Intriguingly, these navigation actions took the form of temporal scan patterns spanning the abstraction levels. This suggests that *functional integration*, that is, integration between the abstraction levels is critical display dimension that should be supported when designing large ecological displays.

This work was funded by a grant from ABB Corporate Research, Heidelberg, Germany, Klaus Zinser, Grant Monitor. This work also received support through an NSERC post-graduate fellowship to the author and an NSERC research grant to Kim Vicente. I would also like to thank Kim Vicente for his guidance of this research.

References

BURNS, C. M. (in press). *The effects of spatial and temporal proximity of means-end information in ecological display design for an industrial simulation*. Unpublished Doctoral Dissertation, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada.

BURNS, C. M. (in press). *Putting it all together: improving display integration in ecological displays.*

JANZEN, M. J. & VICENTE, K. J. (1998). Attention allocation within the abstraction hierarchy. *International Journal of Human–Computer Studies*, **48,** 521–545.

MORAY, N. (1986). Monitoring behaviour and supervisory control. In K. BOFF, L. KAUFMAN & J. THOMAS, Eds. *Handbook of Human Perception and Performance*, pp. 40-1–40-51. New York: Wiley.

PAWLAK, W. S. & VICENTE, K. J. (1996). Inducing effective operator control through ecological interface design. *International Journal of Human–Computer Studies*, **44,** 653–688.

RASMUSSEN, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on Systems, Man and Cybernetics*, **15,** 234–243.

SARTER, N. B. & WOODS, D. D. (1997). Team play with a powerful and independent agent: operational experiences and automation surprises on the Airbus A-320. *Human Factors*, **39,** 553–569.

TORENVLIET, G. L., JAMIESON, G. A. & VICENTE, K. J. (1998). Making the most of ecological interface design: the role of cognitive style. *Proceedings of the 4th Symposium on Human-Interaction in Complex Systems*. Piscataway, NJ: IEEE.

VICENTE, K. J. (1992). Memory recall in a process control system: a measure of expertise and display effectiveness. *Memory & Cognition*, **20,** 356–373.

VICENTE, K. J., CHRISTOFFERSEN, K. & PEREKLITA, A. (1995). Supporting operator problem solving through ecological interface design. *IEEE Transactions on Systems, Man, and Cybernetics*, **25,** 529–545.

VICENTE, K. J., MUMAW, R. J. & ROTH, E. M. (1997). *Cognitive functioning of control room operators: final phase*. CEL 97-01. Cognitive Engineering Laboratory, University of Toronto, Toronto, Canada.

WOODS, D. D. (1984). Visual momentum: a concept to improve the cognitive couple of person and computer. *International Journal of Man–Machine Studies*, **21,** 229–244.