Design and Applications of GLANCE: GLanceable Alarm Notification for a User Centered Experience

Laura Tarantino 1,*, Daniela Angelucci 2, Alessandra Bonomo 3, Annalisa Cardinali 4 and Stefania Di Paolo 5

Abstract: The trade-off between awareness and interruption is a crucial aspect in network fault notifiers: Low severity alarms should not distract operators from other primary tasks, however it might be crucial that operators promptly react to critical notifications. A notification system should hence determine when a particular interruption is appropriate and how it should be presented. In this direction, this paper presents a multistep design path beginning from the objective of designing a proof-of-concept for a glanceable alarm notification component for telecommunication network management systems based on a peripheral display approach. In particular the goal was a notifier guided by severity-based strategies and offering the information expressiveness of a one-notification-at-the-time perspective while enriching it with overview capabilities to guarantee (possibly subliminal) long-term local and global content comprehension and prompt reaction only when the interruption from the foreground task is dictated by the fault severity. A first design macro-phase led to the simple yet effective GLANCE (GLanceable Alarm Notification for a User Centered Experience) model, based on a visual coding technique oriented to comprehension and reaction, and a transition strategy oriented to interruptions and reaction. A second design macro-phase studied the application of GLANCE to a personal customizable multichannel notification tool and to a service-oriented fault monitor for digital terrestrial television broadcasting networks.

Keywords: information visualization; notification systems; peripheral displays

1. Introduction

Network management refers to activities, methods, procedures, and tools that pertain to the operation, administration, maintenance, and provisioning of networked systems [1]. In particular, fault management pertains to the notification, isolation, and correction of, as well as revealing, network faults. A variety of network monitoring tools are available on the market offering, to network maintenance operators, suites of software packages that provide, among others, automated failure detection and support for controlling multiple devices through graphical user interfaces (see, e.g., [2–4]). Generally, these platforms are desktop or web applications based on a WIMP (Windows Icon Menu Pointer) paradigm, with an interaction environment characterized by a multiple view approach that gives the operators the possibility to inspect the network at different levels of details and/or under different perspectives. While successful when used as primary tasks by operators focused on network analysis, they turn out to be less effective as fault notifiers for operators attending to other primary tasks (possibly in a mobile context) and that have to be informed of faults in network elements: In-desktop fault notification often lacks proper mechanisms to capture the correct degree of an operator’s attention based on the alarm severity, so
to interrupt the primary task only when necessary at the right moment, while mobility-oriented fault notification is often based on email or SMS, forcing operators to interrupt their activities to read the message, check the alarm severity, and access (textual) information about the alarm description.

The cost-benefit trade-off between awareness and interruption is a crucial aspect [5]: Interruptions unrelated to the primary task may negatively influence task performance and affective state, leading to longer task completion times, higher error rates, and increased frustration [5–8] and, on the other hand, studies have proven that when users are offered the possibility of negotiating the receipt of notifications they tend to an indefinite postponement [9], with a resulting inability to get the right information at the right time [7,10,11], which would be highly harmful when the recipient of the notification is expected to be responsible for some relevant action in response. For example, the study discussed in [12] shows how, in a clinical environment, sending digital notifications with reminders of incoming appointments to patients greatly improves the efficiency of health care services by increasing attendance and sensibly reducing the number of “no shows”. Similarly, in contexts like network management, of interest for this paper, task switching in response to a fault alert may be mandatory when the operator must promptly react and take charge of a critical request, while being inopportune in case of low severity faults. Furthermore, since more than one alarm can be “active” (i.e., raised and not acknowledged yet) in the network in a given moment, it is advisable that network management notifiers be able to also offer a global overview of all incoming active alarms. In the remainder of this section we provide a brief overview on some issues affecting the design of notification delivery, singling out the motivations behind the proposed study, and outlining research questions and the overall structure of the study.

1.1. Related Literature

Notifications can be defined as proactive cues based on (a combination of) visual, auditory, and/or haptic alerts delivering information outside the current user’s focus of attention (e.g., [13–15]). Depending on the kind of static and dynamic stimuli adopted, a notification may behave as a trigger to switch from a primary to a secondary task, with the risk of consequent disruption and user frustration. The design of notification delivery must hence be based on balancing the necessity of keeping the user informed about incoming information and the necessity of minimizing the pervasiveness and invasiveness of notifications (sometimes defined as “interruptiveness”). Literature studies on notifications are primarily conducted along two main research lines:

- Studies on the impact that notifications delivered on single or multiple devices have on a variety of interrelated aspects: Users’ performances in primary tasks interrupted by notifications, users’ awareness on secondary tasks gained through notifications, and users’ frustrations due to task switching;
- Studies on architectural aspects and infrastructures for notification delivery on single or multiple devices are becoming increasingly relevant along with the increasing availability of possibly coordinated connected devices (e.g., [16–20]).

As to the first research line, a great volume of psychology studies on task interruption was stimulated since the 1990s by the growing diffusion of digital communications, revealing the intricacy of related design decisions (see, e.g., [13,21,22] for overviews on these studies). None of the interruptions policies (immediate, negotiated, mediated, or scheduled) singled out in [9] proved to be the definitive (and most effective) response to the problem of interrupting users while guaranteeing minimal effects on their performance. A number of studies investigated the effect of notifications and interruptions on users of desktop PC, especially in work environments [21–27]. In particular, the study in [21] was conducted to evaluate the influence of Instant Messaging on memory and performances of the primary task (web search tasks were used in the experiment as primary tasks), showing that the harmful effects of the interruption vary depending on the time of arrival relative to the task’s lifecycle (the earlier the stage of the task, the worse the effect). The week-long in
situ diary study in [22] was conducted to characterize how people define tasks, and with which granularity, how they interleave multiple tasks among interruptions, and the level of details with which they encode “task switch”. In particular, the study singled out the factors influencing the perceived difficulty of switching back to tasks and showed that reacquiring “returned-to” tasks is considered difficult by users. With the aim of reducing frustrations due to the inopportune moments in which interruptions arrive, the study in [23], based on statistical models, evaluated how the adoption of a notification management system scheduling notification delivery would impact PC users and their tasks with the results showing that users’ frustration indeed decreases when adopting defer-to-breakpoint notification policies (hence requiring the detection of adequate breakpoints during interaction tasks). A different approach to the mitigation of “interruptiveness” can be found in [24,25], which evaluated to which extent social strategies (e.g., the use of empathy or of politeness in notification messages) could reduce the negative effects of interruptions. These studies showed that a polite interruption applied in a low cognitive load condition may indeed reduce “interruptiveness” but that, on the other hand, empathy can be counter-productive while interrupting a high cognitive load task for being perceived superfluous and more irritating than a neutral message. Disruptions can be even heavier in mobile contexts, with notifications crossing the line between work and private life, even if the variety of contexts and the notification nature make the analysis more complex than in work settings. For example, the one-week in situ study in [13], involving 15 users, confirmed the negative emotions associated, in general, to a high volume of notifications, but showed also that deferring strategies may work well in professional settings while notifications associated to personal communication increased positive connection feelings (deferring strategies may also be difficult in mobile contexts with uncontrolled naturalistic settings and unpredictable primary tasks [28]). As to the impact of the notification nature, the in-the-large assessment in [14], collecting about 200 million mobile phone notifications from about 40,000 users, studied the difference of the perceived importance of various types of notifications. Similarly, other studies recognized that notifications are used as an awareness mechanism: A field study about email communication [15], conducted with PC 20 users, revealed that, while aware of the distractions caused by notifications, users highly valued the awareness of incoming information provided by notifications and used this to decide whether to suspend ongoing tasks and switch to the source of the notification. Furthermore, one may notice that, in some multitasking application domains, external interruptions are proved to be even less disruptive than self-interruptions (i.e., voluntary task-switching) [29].

1.2. Dealing with Notifications in Network Management

In summary, the basic approach underlying the studies about “interruptiveness” is to reason on the timing (the when) of the interruption, under the hypothesis that the notification delivery can be subject to some sort of scheduling based on the stages of the ongoing tasks (in other words, the ongoing task “wins” over the source of the notification). As already observed, in an application domain like network management, this is not always the case, with the “winner” between the ongoing task and (source of the) notification being dependent on the severity of the fault.

An alternative solution can be found by reasoning about the format (the how) of the notification, making it dependent on the notification severity and capable of capturing the correct degree of users’ attention: While the customary solution (found both on desktop PCs and in mobile devices) is to adopt a unified format for all notifications (typically a badge possibly associated to auditory and haptic cues), a severity-based solution would associate low severity alarms to notification formats that do not distract operators from the foreground tasks and high severity alarms to notification formats that provoke the operator’s reaction. This is the approach proposed in this paper, based on a glanceable peripheral display, that, residing in the sensory periphery, conveys information that can be of some importance for users, in a way that it is able to exploit the “pre-attentive” processing
ability of the human brain; as long as the alarm severity is not high, the display behaves like an ambient display [30], which lets users absorb the information in a subliminal way, without having their attention diverted from the foreground task (ambient systems are diffused in a variety of contexts, like, e.g., [31–33]). Only when a high severity notification arrives, the display changes its (static and dynamic) features to behave like an alerting display moving from the periphery to the foreground of the user’s attention.

Another aspect of customary solutions that make them poorly performant in the case of network management is their perspective of streams of individual notifications displayed one at the time and inserted in lists ordered by the time of arrival without any overview capability. The burden of creating a mental map of the overall state (who/which source sent what when) is entirely on the users that can only scroll lists of badges. A notifier oriented to network management should be capable of offering, at the same time, information on individual recent faults and overview on the network as a whole. The one-notification-at-the-time perspective is generally also typical of studies focusing on features and classifications of peripheral displays, design guidelines for criticality-based displays, and the prioritization of notifications [34–41]. One notable exception is the Scope notifier [42], offering an overview on multiple prioritized notifications coming from heterogeneous sources (among which, email, calendar, and pending work items) using the metaphor of a traditional radar view (the closer to the center, the most urgent/important the item). The Scope peripheral awareness mode is extremely abstract and concise, representing notifications just as markers on the radar and leaving the initiative primarily to the users that are required to switch from the foreground task to the Scope active interaction mode to inspect items and decide whether to open it in the native application (only for urgent notifications details are automatically displayed). On the contrary, we aim at a notifier that always retains the information expressiveness of the one-notification-at-the-time perspective, though guided by severity-based coding strategies, while enriching it with overview capabilities, to guarantee (possibly subliminal) long-term local and global content comprehension and prompt reaction when the interruption from the foreground task is dictated by the fault severity.

In this direction, this paper describes research that originated within the framework of cooperation between the University of L’Aquila and Technolabs S.p.a, a former Italian ICT company located in L’Aquila, and aimed at designing a proof-of-concept for a glanceable alarm notification component for Telecommunication (TLC) network management systems, based on a peripheral display approach. The adopted methodological approach allowed us to maintain focus on the company’s problem-solving interest and research interest, favoring the generalization of achieved results. The study was conducted with the following main research questions in mind: “Can the fault notification component of a network monitoring tool be effectively defined according to a peripheral display paradigm?”, and more specifically “To which extent is it possible to offer a glanceable overview along with information on individual faults?” and “Can (and to which extent) results achieved for a TLC network be generalizable to a more general context of information coming from heterogeneous sources?” This first design macro-phase led to the simple yet effective GLANCE (GLanceable Alarm Notification for a user Centered Experience) model, based on a visual coding technique oriented for comprehension and reaction, and a transition strategy oriented to interruptions and reaction. A second design macro-phase studied the applicability of GLANCE results to a personal customizable multichannel notification tool and to a service-oriented fault monitor for digital terrestrial television broadcasting networks, also leading to an enrichment of the model.

The remainder of the paper is organized as follows. Section 2 provides an overview of the two design macro-phases underlying overall and intermediate objectives. Then Section 3 discusses results from the first design macro-phase and Section 4 surveys the results of the application studies of the second design macro-phase. Section 5 discusses the overall design experience and its results, and, finally, Section 6 draws conclusions.
2. Method

The development of GLANCE followed an evolutionary path. Though the initial goal was quite specific (focused on a system for optical fiber TLC networks in a specific company), the analysis of the involved critical information led us to a notifier exhibiting a level of abstraction and a generality that stimulated the investigation of its applicability in different contexts as well. The overall development hence proceeded along two macro-phases, respectively focused on a first definition of the model and on its extension and application. The first macro-phase was carried out in cooperation with Technolabs S.p.A., a former Italian ICT company, while the second macro-phase involved two sequential User Centered Design (UCD) based projects aimed at testing the applicability of the achieved results, the first carried out within the University of L’Aquila, and the second carried out in co-operation with Broadcast Solutions S.R.L., an Italian company offering system integration capabilities for broadcasting and media industries.

2.1. First Macro-Phase: The GLANCE Project

Given the twofold objective of (1) providing a solution to a specific company problem and (2) generalizing technical results to achieve research results, the structure of the general organization of the first macro-phase was inspired by the “tandem model” proposed by [43] within the realm of Action Research (an iterative methodological approach based on the juxtaposition of action/practice and research/theory [44,45]). The general idea is that to enforce the mission of production of scholarly knowledge while seeking solutions for a real-life practical problem proposed by an organizational “client”, two cycles run in parallel (see Figure 1), one addressing the client’s problem-solving interest and the other one addressing the researchers’ scholarly interest, leading to the desired results after a number of iterations [45,46] (in Action Research the most utilized structure for a single iteration is the one proposed by [47], including the five stages of diagnosis, action planning, action taking, evaluation, and learning, conducted under theoretical premises singled out during the diagnosis stage of the first iteration). As underlined by [48], Action Research can be successfully and creatively employed in empirical Human Computer Interaction (HCI) research, and has received attention by HCI researchers (see, e.g., [48–53]). Within the framework of the GLANCE project, we borrowed from the tandem model the ideas of (a) separating yet running in two parallel cycles, the research objective and problem-solving objective and (b) treating the outcome of the research cycle (iteratively evaluated and revised) as the solution method of the problem-solving cycle.

![Figure 1. The tandem model.](image-url)

The research started within the framework of ongoing cooperation between the University of L’Aquila (in the following denoted as Univaq) and Technolabs S.p.A, with the specific goal (the problem-solving interest) of enhancing the notification of alarms aimed at...
soliciting technical intervention in a specific optical fiber TLC network. The investigated organizational context involves alarm information that may be classified along two dimensions:

- **Alarm class**: Communication, environment, equipment, processing, quality, and security), and
- **Severity level**: Cleared, warning, minor, major, and critical are the five standard levels in this application domain.

Additionally, each individual alarm is associated to a tuple of values (referring to temporal, network, object, cause, and location information) and has the life cycle illustrated in Figure 2. The “alarm raised” and “alarm cleared” transitions are notified by the system to all network operators, while the “acknowledge” transition is communicated to the system by the operator taking charge of the technical intervention. Alarms remain “active” until acknowledged by an operator. Multiple alarms may be simultaneously active because of faults generated in the same Network Element (NE) or in distinct NEs: With respect to notification purposes, we consider simultaneous alarms “active at the same time” (i.e., for which no operator acknowledgement has been received yet), even if it has not necessarily “arrived at the same time”. It is hence desirable that an alarm notifier be able to provide information not only about a single incoming alarm but also about all active alarms through a global overview.

![Figure 2. State diagram of an alarm.](image)

As pointed out in the Introduction, commercially available network monitoring tools may not prove to be effective as notifiers for operators attending to other tasks. After the analysis of virtues and flaws of traditional systems, and after the context analysis, Univaq and Technolabs researchers agreed on studying and evaluating a novel fault notification approach based on a peripheral display providing severity-based notifications of incoming alarms including glanceable information on both individual alarms and network overview, small enough to be possibly visualized on (possibly ad-hoc hand-held or wearable) mobile devices. This reflection phase was actually the entrance of the cyclical process conducted under the theoretical premise that a shift of paradigm towards peripheral displays could improve operators’ work. The working group conducting the research included three people from Univaq (one senior researcher and two junior researchers with a background in HCI) and two people from Technolabs (one senior electronic engineer with a background in TLC network management and one senior electronic engineer with a background in TLC network management and Information Visualization). Technolabs represented the problem-solving interest, while Univaq represented the research interest. It was agreed that the Univaq researchers had to guide the overall process while the Technolabs people had to be involved as domain experts/stakeholders and to assess design choices. In summary, the elements of the GLANCE project can be summarized as in Table 1.
Table 1. Elements of the GLANCE (GLanceable Alarm Notification for a User Centered Experience) project: A is the investigated real-world problem situation, P is a real-world example of A allowing the researcher to investigate A, F is the theoretical premise declared by the researcher prior to the intervention in A, M\(_R\) is the research method, and M\(_PS\) is the method employed to guide the problem solving (PS) intervention.

| Elements of the GLANCE Project |
|-------------------------------|
| A | Issues/challenges in handling simultaneous notifications with different severity levels |
| P | Setting up a proof-of-concept notifier for faults in optical fiber networks |
| F | Glanceable peripheral displays may improve the efficiency of operators’ work |
| M\(_R\) | HCI methods/techniques blended with organization elements from the “tandem model” |
| M\(_PS\) | The novel notification model (GLANCE) |

As to project and research objectives, though presented hereinafter all together for the sake of a clear presentation, they were actually the result of continuous reflection and refinements throughout the iterations of the cyclical process. In order to investigate the three identified general research questions, Q0 “Can the fault notification component of a network monitoring tool be defined according to a peripheral display paradigm?”, Q1 “To which extent is it possible to offer a glanceable overview along with information on individual faults?” and Q2 “Can results achieved for TLC network be generalizable to a more general context of information coming from heterogenous sources?”, the problem objective (PO) and the research objective (RO) for the two parallel cycles, along with their intermediate goals, were defined as in Table 2.

Table 2. Problem and research objectives.

| Problem Objective (PO) and Intermediate Goals |
|-----------------------------------------------|
| PO Setting up of a proof-of-concept severity-based multiple fault notifier for Telecommunication (TLC) networks, with intermediate goals: |
| PO1 Notification of incoming single alarms |
| PO2 Notification/overview of multiple active alarms |

| Research Objective (RO) and Intermediate Goals |
|-----------------------------------------------|
| RO Definition of a general abstract model for the notification of multiple alerts coming from heterogenous sources |
| RO1 Definition of an abstract model for the notification of single faults in TLC networks |
| RO1.1 Definition of an abstract display schema oriented to comprehension and reaction for the visualization of the information associated to a single alarm |
| RO1.2 Definition of a dynamic stimuli strategy oriented to interruption and reaction in response to the notification of a single alarm |
| RO2 Definition of an abstract model for notification/overview of multiple active faults in TLC networks |
| RO2.1 Extension of RO1 results to the case of intra-NE (Network Element) multiple active alarms |
| RO2.2 Extension of RO1 results to the case of inter-NE multiple active alarms |
| RO2.3 Extension of RO1 results to the combined case of multiple active alarms |
| RO2.4 Definition of constraints for the number of multiple active alarms |
| RO3 Release of a general abstract model for the notification of multiple alerts from heterogenous sources |

According to the specification of problem and research objectives, it was agreed that GLANCE had to be developed through an evolutionary design approach starting from the simplest scenario in which we have only one active alarm, successively moving to the scenario with multiple active alarms to be separately addressed in the first two iterations of the cyclical process (on both the research and problem-solving interest cycles), and a third iteration would then be focused on the generalizability of results (mostly on the research interest cycle), as summarized in Table 3.
Table 3. Overall project plan.

| Iteration | Main Objective | Evaluation |
|-----------|----------------|------------|
| 1st iteration | dealing with individual alarms | expert based |
| 2nd iteration | dealing with multiple alarms | expert based + user based + dissemination |
| 3rd iteration | generalization of results and definition of the GLANCE model | expert-based + user based |

Formative usability evaluation methods [54,55] have been used throughout the whole design process to iteratively assess design choices, not only with the objective of detecting and eliminating usability flaws but also with the aim of informing the design based on objective measures where necessary (e.g., to provide an answer to research question Q1 “To which extent is it possible to offer a glanceable overview along with information on individual faults?”). All the experiments mentioned in the paper are customary usability evaluation, conducted with adult people within the framework of official Master thesis works of the School of “Ingegneria Informatica e Automatica” of the University of L’Aquila (Italy) approved by the institutional School Review Board [56–59].

2.2. Second Macro-Phase: GLANCE Applications and Interactivity Enrichment

While the company interest was considered to be almost satisfied at the end of the second iteration, the research interest could be regarded only partially satisfied at the end of the third iteration: Formal aspects were indeed judged sound by expert-based and user-based usability evaluation studies, however, given the ambition of a model with general applicability, it was considered appropriate to conduct additional research to test the model’s effectiveness and flexibility in contexts that are diverse in nature, so as to be used as benchmarks. In particular, two traditional UCD-based projects were focused to:

• The application of GLANCE to a personal customizable multichannel notification tool;
• The application of GLANCE to a service-oriented fault monitor for digital terrestrial television broadcasting networks.

The first application led to GLANCE enrichment in terms of both visual coding and (new) interactive features. The revised model (GLANCE+) turned out to be also beneficial for the second application project.

The overall evolution of GLANCE features (visual coding technique, transition strategy, and interactivity) throughout the design project, along with their evaluation, is discussed in the next two sections, which are focused on the two macro-phases of the project.

3. Results of the 1st Macro-Phase: The GLANCE Project

As we already observed in the Introduction, our first objective was to overcome the limits that customary approaches have with respect to requirements of network management, in particular we wanted to abandon the perspective of streams of notifications displayed one after the other through badges with a unified format and inserted in a scrollable list based on timestamps, to achieve a notifier acting as an ambient or an alerting display based on severity-based coding strategies, retaining the information expressiveness of the one-notification-at-the-time perspective, and enriching it with overview capabilities. The objective is to guarantee (possibly subliminal) long-term local and global content comprehension and prompt reaction when the interruption from the foreground task is dictated by the fault severity (the three parameters Interruption (I), Reaction (R), and Comprehension (C) have been proposed by [36] as pillars of the so-called IRC framework). Context-of-use analysis, user observation, field studies, and brainstorming meetings with domain experts in the initial phases led to the founding idea of a system responsible for selecting a Level-Of-Detail (LOD) and dynamic stimuli consistent with the urgency degree of the notification, based on a design “serving” the IRC parameters by (1) a visual coding technique (which is the leading aspect of the study) oriented to comprehension and reaction and (2) a transition strategy, oriented to interruptions and reaction.
In this section we discuss the design of GLANCE, addressing separately the two cases of single alarm and multiple alarms, faced from both the research and problem-solving point-of-view in the first two iterations (in Sections 3.1 and 3.2, respectively), and then discussing the generalizability of results (in Section 3.3).

3.1. First Iteration: Dealing with Individual Alarms

The first iteration was aimed at the achievement of the research objective RO1 (Definition of an abstract model for the notification of single faults in TLC networks) and of the problem objective PO1 (Notification of incoming single alarms).

3.1.1. Definition of the Abstract Model and Application to the Organizational Situation

The basic idea was to conceive the notifier as an in-desktop peripheral display located in the bottom-right corner of the primary monitor, outside the primary visual focus, occupying a fixed rectangular area small enough to be also possibly displayed on hand-held mobile devices or on ad-hoc (wearable) devices (studies have shown that small displays result in fast identification of changing information [37]). By default, the display behaves like an ambient display, continuously visualizing non-critical information without demanding the full attention of users, but as soon as a critical notification arrives, it turns to an alerting display. In particular:

- Low severity alarms are associated with a few data conveyed in a subliminal way, whereas;
- More urgent alarms are associated with notifications requiring focal attention and technical intervention,

by adapting the visual coding technique and the transition strategy to the severity level. Given their distinct responsibilities with respect to the IRC framework, the definition of these two components were addressed separately. In both cases the approach was to define general abstract rules and then apply them to the specific problem at hand.

The GLANCE visual coding technique. As observed in the Introduction, one of the basic choices behind GLANCE is to abandon a unified notification format (adopted, e.g., in customary notification badges in PC desktops or smartphones) in favor of an adaptive format, taking into account the severity degree of the notification. The final objective is a glanceable display able to automatically convey quickly understandable but minimum and specific information (with respect to the severity) without requiring any user action. In order to ensure glanceability, the visual coding technique is based on visual variables (color, shape, size, position, and contrast). More specifically, the basic GLANCE general abstract schema is defined as a rectangular display partitioned into three sub-areas (see Figure 3):

- An upper bar contains temporal information on the displayed alarm,
- A lower bar is reserved to user data, and
- A middle area contains information about the currently active alarm and is in turn divided into a synthetic component and a detailed component: The former informs about the alarm severity by a combination of textual and sensorial information (label and color), while the latter visualizes the alarm description with the background color and the LOD associated to the severity, conveying the minimum amount of information necessary for the severity level and getting progressively denser information while the alarm severity increases.

Application to the specific problem situation. As mentioned in Section 2.2, each alarm is described by a tuple of data item. As to LOD, three levels were judged appropriate, determining a ranking of the attributes of the tuple schema, as illustrated by Table 4. Alarms with a different severity level and same LOD will be differentiated by the values of retinal variables and the transition types.
Figure 3. First GLANCE general abstract schema: (a) The three components and (b) how alarm information is mapped onto the synthetic and detailed component.

**Table 4.** Level-Of-Detail (LOD) defined as association between severity level and schema attributes.

| LOD  | Severity Level | Schema Attributes                          |
|------|----------------|-------------------------------------------|
| LOD1 | cleared, warning | temporal data, network, NE id, cause      |
| LOD2 | minor, major    | temporal data, network, NE id, class, object, cause |
| LOD3 | critical        | temporal data, network, NE id, class, object, cause, location, sub-location |

Figure 4a shows how the five levels of severity are mapped onto three abstract schemata according to the LOD conventions of Table 4. In the case of critical alarms, the quantity of data items associated to the corresponding LOD3 (cause, location, and sublocation) requires the utilization of a string ticker with the beneficial side effect that a dynamic stimulus-like ticker animation will grab users’ attention coherently with a high severity alarm (additional dynamic stimuli will be determined by the transition strategy, discussed later on). Examples of alarm visualizations with different severity are depicted in Figure 4b, illustrating the roles of icons to represent alarm class and background color to represent severity (it is worth recalling that results on glanceability tests indicate that these two visual properties are the two most popular among users [39]). In order to not violate user expectations, color coding conforms to the standard severity/color association used in the network management application domain (cleared/green, warning/blue, minor/yellow, major/orange, and critical/red). It has to be observed that in any case additional coding mechanisms (textual, auditory, and possibly haptic cues) are used in association to colors to guarantee a correct display interpretation to users with color vision deficiencies: The synthetic component also displays the severity level as a textual label, the detailed component conforms to different LODs depending on the severity level (e.g., no icons are used for LOD1), and distinct dynamic (animation, auditory, and haptic) stimuli are used to grab the correct level of user’s attention depending on the severity level.

**GLANCE transition strategy.** The basic idea behind GLANCE is that low severity notifications have to remain in a pre-attentive stage of the human brain, while high severity notifications must interrupt ongoing activities. The second component of GLANCE is hence an attention capture mechanism used to determine how the notification display shifts from ambient to alerting, i.e., from visual periphery to foveal vision, and vice versa, and how to control awareness, and later processing. Since the display is located in the visual periphery, such a mechanism has to be based on dynamic stimuli (e.g., animation) depending on the severity level and stages of human attention. Studies show that human attention may be divided into three main zones [38]: In the zone of inattention, objects may affect behavior though not directly available for conscious awareness; in the zone of divided attention, where users’ attention is distributed over a number of distinct objects; and finally, in the zone of focused attention, where all users’ attentional resources are concentrated onto only one stimulus (Figure 5). Based on this attention model, it is straightforward to consider
three notification levels (low, medium, and high), and correspondingly three transition types, associated to the three zones of human attention. It is also advisable to foresee a fourth notification level in which not only users reach the zone of focused attention, but they also perform some acknowledgement action. In GLANCE we adopt the transition framework proposed and validated in [38] (Table 5), where the four levels are denoted change-blind (information of marginal importance to be notified so to consume no conscious awareness), make-aware (information directed to divided attention, of which user has to be made aware), interrupt (information directed to focal attention, demanding interruption from current tasks), and demand-attention (notification with the highest urgency that has to distract the user from the primary task until an action is done). The third column in Table 5 shows how display updates are associated with visual/auditory stimuli of different types based on consolidated literature results (surveyed and structured in [38]) guaranteeing to the notifier the correct behavior of change-blind/ambient/alerting system based on the severity level (in case of usage on mobile devices, distinct auditory and haptic stimuli of increasing intensity are to be considered for each level).

![Figure 4](image-url)

**Figure 4.** Application of the first GLANCE model to fault notification: (a) Mapping the five severity levels onto three abstract schemata and (b) sample fault notifications with different severity levels.

_Application to the specific problem situation._ The five severity levels considered in the TLC network naturally maps to different areas of the attention-awareness graph, as illustrated in Figure 5: Cleared and warning are associated with the inattention zone and the change-blind notification level, minor is associated with the divided attention zone and the make-aware notification level, major is associated with the focused attention zone and the interrupt notification level, and critical is associated with the focused attention zone and the demand-attention notification level.
In case of usage on mobile devices, distinct auditory and haptic stimuli of increasing intensity are to be added.

3.1.2. Evaluation and Learning

The first (low-fidelity) prototype was realized as an animated PowerPoint display mock-up simulating an example sequence of single alarm notifications with different severity. Given the simplicity of the notifier up to this point, formative evaluation was conducted in a rather quick and inexpensive way to receive suggestions from domain experts on how to possibly improve the glanceability and readability of the system, postponing a
more formal evaluation to the end of the 2nd iteration. The evaluation group included one facilitator, one product expert, and two Technolabs people with the twofold role of evaluators and domain experts. The overall evaluation meeting included a 30′ seminar on ambient and alerting displays (to make domain experts more familiar with the selected theoretical framework), a general presentation of the GLANCE model to the evaluators, and a cognitive walkthrough session [60,61].

In order for the test to be as much as possible representative of an actual working situation (e.g., to check how typical background noise could interfere with the auditory stimuli of the notifier), the evaluation session took place in the real working setting and the notification sequence was built from a real alarm log generated by the optical fiber network investigated in the problem solving cycle, with notifications presented in an asynchronous way with no particular order as to severity levels. The facilitator walked through the example notification sequence to evaluate, for each notification, whether a real user would achieve the desired comprehension effect and would associate it to the correct reaction. Technolabs people were indeed very satisfied and surprised by GLANCE readability and did not detect any flaw neither in the abstract model nor in the mockup. In particular, they appreciated the glanceability of the coding technique and underlined the immediacy of the alarm information availability compared to solutions offered by commercial products based on a peripheral status bar with scarce visual cue distinction and, above all, on the necessity of user initiated actions for accessing the correct level of detail for incoming alarms.

In addition, stimulated by this observation, after the walkthrough session a comparison between GLANCE and the Scope display [42] was analyzed with the domain experts, as to the different choice of the two approaches with respect to the immediate availability of the information associated to low and medium severity alarms (as to information about high severity notifications both approaches make it immediately available without explicit user request). Domain experts agreed with the GLANCE choice of making alarm information always available at an automatic severity-dependent LOD without explicit user’s action in any case, and, concerning this stage of the project, judged GLANCE superior as to long-term comprehension of alarm information of any severity (though additional evaluation was still to be done in successive project steps).

With reference to Table 1, the target objectives of the first iteration (PO1, RO1.1, and RO1.2, and then RO1 as well) were hence considered achieved.

3.2. Second Iteration: Dealing with Multiple Alarms

The second iteration was aimed at the achievement of research objective RO2 (Definition of an abstract model for notification/overview of multiple active faults in TLC networks) and problem objective PO2 (Proof-of-concept ability for notification/overview of multiple active alarms).

3.2.1. Definition of an Abstract Model for Notification/Overview of Multiple Active Faults

The main objective of this stage was the enrichment of the basic model with an overview capability of supporting the final users in the cognitive task of building a mental map of the network status without having to interact with the device for querying the system: The basic idea behind GLANCE is in fact that network overview and information on individual alarms have to coexist and continuously inform users without requesting any action from them.

As observed in Section 2.1, the problem is to conceive a glanceable synthesis of simultaneously active alarms (i.e., alarms which have not yet been taken in charge by some operators), possibly prioritizing alarms with higher severity. In TLC networks, multiple active faults may be generated within a single NE, within distinct NEs, or in a hybrid situation. To adapt the display to manage these cases yet retain the information expressiveness of the one-notification-at-the-time perspective, we assigned to the synthetic component the role of glanceable overview of the overall situation, to provide quick and easy comprehension of the number of alarms, their severity, and their sources (same NE, distinct NEs,
or both). The detailed component maintains its role of container for alarm descriptions, but it is used differently depending on whether the simultaneous alarms are originating in the same NE or in distinct NEs:

- **Case 1 (intra-NE):** *Simultaneous faults in the same NE* (Figure 6a illustrates an example for the specific problem situation): In this case the synthetic component is split into as many portions as the number of simultaneous alarms, each showing the severity of the associated alarm, ordered by severity level. The detailed component shows the alarm with the highest severity level, representing less critical ones as annotated flags;

- **Case 2 (inter-NE):** *Simultaneous faults in distinct NEs:* In this case the synthetic component is divided into as many portions as the number of NEs generating the faults, with each portion showing in turn the severity of the associated alarm, ordered by severity level. The detailed component cyclically visualizes the descriptions of the incoming alarms one after the other, as a customary sequence of alerts ordered by criticalness (Figure 6b illustrates an example for the specific problem situation, showing a sequence of two active alarms originating in two distinct NEs), in other words, the notifier continuously cycles over alarmed NEs;

![Figure 6. Redefinition of the display format in the case of multiple notifications: (a) Multiple faults in the same NE (Case 1) and (b) sequence of multiple faults in two NEs (Case 2).](image)

- **Case 3:** *Hybrid situations:* In this case the approaches utilized for facing Case 1 and Case 2 are combined, with the synthetic component split into as many portions as the number of NEs generating the faults, like in Case 2. Each NE is assigned the same area but this area is further partitioned when the NE generates more than one fault (consistently with Case 1). As in Case 2, the detailed component visualizes alarm descriptions as a sequence of alerts, continuously cycling over alarmed NEs (for multiple faults coming from the same NE, only the most severe one is displayed, consistently with rules for Case 1).

### 3.2.2. Evaluation and Learning

The second (low-fidelity) prototype was realized as an animated PowerPoint display mock-up simulating an example sequence of single and multiple alarm notifications including a variety of intra-NE, inter-NE, and hybrid case scenarios. This round of formative evaluation included (1) an expert-based evaluation session with Technolabs people followed by (2) a user-based test aimed at evaluating the readability of the display and at measuring the scalability of the technique.

#### The Expert-Based Evaluation

The evaluation meeting was organized and conducted similarly to the evaluation session performed at the end of the 1st iteration, with the goal of evaluating *glanceability, readability, and learnability* of the notifier in the new context of the added overview capability. The meeting included a presentation of the enriched model, a cognitive walkthrough session with domain experts, an informal think aloud testing with experienced users,
and a final discussion with both domain experts and experienced users. Overall it lasted two and a half hours. The evaluation group included one facilitator, one product expert, two Technolabs people with the twofold role of evaluators and domain experts (who also participated in the previous evaluation session), and two Technolabs people with the role of experienced users (who did not participate in the previous evaluation session). Again, the session was conducted in a real working environment and the examined notification sequence was built from a real alarm log generated by the optical fiber network under consideration in the problem-solving cycle. In particular, the sequence was set up by collecting a number of log portions selected so to include progressively different combinations of intra-NE, inter-NE, and hybrid case scenarios, with notifications presented in an asynchronous way. The same sequence was used for the cognitive walkthrough session with domain experts and the think aloud testing with experienced users. In both cases, the facilitator walked through the example notification sequence to evaluate whether real users would achieve the desired comprehension effect and would associate the correct reaction. In particular, in the think aloud testing users were asked to comment on how they interpreted the visualized information so as to evaluate the learnability of the GLANCE display and its readability against a benchmark situation, in particular in the tricky case of hybrid situations (Case 3), with the main attention on the comprehension of the network status based on the overview provided by the synthetic component. Experienced users involved in the think aloud testing did not participate in the walkthrough session, so to obtain some first rough indications on the learnability of the system: The test was indeed promising, with users grasping correctly the network status from the status of the synthetic component after few steps of the sequence, necessary to understand that the display was cycling over alarmed network elements.

Beside the legibility of the synthetic component, the other crucial aspect to be evaluated and discussed was the necessary visual representation power trade-off (dictated by space constraints) between overview information and individual alarm information, and in particular, the choice of reserving the detailed component to one single alarm (the one with highest severity) for each alarmed network element. All domain experts and users judged the choice appropriate with reference to the main goal of a notification in this context, i.e., \textit{pushing operators to take charge of incoming alarms starting from the most severe ones}, as soon as the most sever alarm in a NE is acknowledged by an operator, it will disappear from the GLANCE display while the next in the priority list of its NE will appear, making the notifier behave like an intervention task dispatcher. It was observed that in an operative release of the notifier, some kind of interactivity would be desirable to allow operators to manually access detailed information on less severe alarms on the same NE on-demand if judged necessary (e.g., the overview might reveal that less severe alarms of a given NE are more severe than displayed alarms of other NEs and operators might wish to inspect all alarms of that NE). Furthermore, the evaluation revealed usability flaws concerning the utilization of flags for the visualization of multiple alarms in the same NE, due to both possible color overlapping (readability flaw) and selection difficulty should the flags become the interactive elements, allowing to access individual alarm data (accessibility flaw).

Furthermore, the GLANCE approach was again compared with the approach of the Scope display, notwithstanding the greater overview capacity offered by the abstractness and conciseness of Scope, in this case domain experts judged GLANCE more appropriate with respect to the application domain, observing that, in the tradeoff between overview and details, sacrificing some overview capacity is absolutely well outbalanced by the advantage of the immediate availability of alarm data in change-blind, ambient, and alerting modes based on the status of the network. It was also appreciated that GLANCE preserves the visibility of the overview in all cases, while in Scope the overview is hidden when a high severity notification arrives. Despite the detected usability flaw and observation about possible interactive capability, Technolabs people were so satisfied with the overall results that a half-day company workshop was organized to share project results with company managers and company laboratory staff. The workshop included a seminar on
ambient and alerting systems and peripheral displays, the presentation of the GLANCE model, a demo of the developed proof-of-concept based on the same notification sequence used for the evaluation session, and a final brainstorming with workshop participants. About 40 people participated in the meeting and gave feedback. Company people reactions to the proof-of-concept demo provided additional evaluation elements, with a clear confirmation of the notifier suitability at provoking reactions correctly depending on the alarm severity. We noticed that although at this stage of the study the association of Table 5 (mapping severity levels to consolidated attention types/transition types) was still to be formally evaluated, it was somehow expected that users’ reactions would be correct, being dynamic stimuli based on state-of-the-art consolidated results [38]. A number of company employees, in particular, underlined the reaction urgency provoked by the dynamic stimuli associated to high severity alarms (which is actually one of the main objectives for the notifier component of a TLC network management system) while many appreciated the “politeness” of the display in the case of low-to-medium severity alarms.

The User-Based Evaluation to Identify the Overview Readability Threshold

As to the user-based evaluation, prior to a test specifically devoted to validate the mapping in Table 5, it was crucial to measure the glanceability and scalability of the solution singled out for overview, so as to provide an objective answer to research question R1 (“To which extent is it possible to offer a glanceable overview along with information on individual faults?”) and inform successive steps of the study. It has to be observed that the decision to assign to the synthetic component the role of global overview makes it the most critical part of the display. In fact, if the widths of individual areas in the synthetic component become too small, the discernibility of individual alarm severity may be jeopardized for two reasons: On the one side, size is a dissociative visual variable that, for low values, affects the perception of colors, and, on the other side, heavy simultaneity increases the probability that contiguous colors do not contrast enough. Furthermore, it may become difficult to grasp in a glanceable way the number of incoming alarms. The proposed solution is hence efficient when the number of simultaneous alarms remains below a given threshold.

Experiment design. To identify the readability threshold, we performed a usability study at a mockup level based on sample configurations of working cases. In particular, we built a series of 40 screenshots covering a broad range of possible scenarios and critical configurations compatible with the display size (e.g., intra-NE simultaneity, inter-NE simultaneity up to six NEs involved, contiguity of very similar or very dissimilar colors) and in this case we used ad-hoc synthetic scenarios in place of sequences extracted from the alarm log to test specific risky combinations. After an internal pilot test aimed at tuning the timing of the administration of screenshots in the sequence, the actual laboratory-based test took place with ten recruited users (an appropriate sample of subjects for the purpose of the test [62].

Experiment sessions. In total 10 (four female) experienced users of PCs with Microsoft Windows aged 22–48 (AVG = 30.8; M = 26) took part in the study on a voluntary basis, recruited among students and personnel of the university campus (five participants were Master students, one participant was a PhD student, two were laboratory technicians, and two were from the administrative personnel). Participants were invited singularly to the lab, welcomed, and informed about the objective of the study. After a brief description of features and behavior of the GLANCE display the test began. Each participant was presented with the screenshots of the sequence, with each screenshot displayed for a few seconds to evaluate what users grasped at a glance (Figure 7 illustrates some of the scenarios shown to test participants). After each screenshot, users were asked to report the number of alarms in the synthetic component for each color (actually, this request is even more demanding than the actual task in a real situation, in which low severity notifications may be ignored with little harm). Overall, each user session lasted approximatively 30 min.
Results. Results are consistent with Miller’s study on the channel capacity and on “the magical number seven, plus or minus two” [63]. In the case of intra-NE simultaneity (Figure 7a) correctness of responses dramatically dropped from above 80% to 40% moving from eight to 10 simultaneous alarms (screenshots were made progressively denser with an increase of two additional subareas in the synthetic component at each new step), while nobody provided a correct answer with 12 simultaneous alarms. The steep discontinuity in performance from the 8 case to the 10 case makes it reasonable to fix 8 as threshold for the case of intra-NE simultaneity (see Table 6). It is worth noticing that since alarm markers are displayed from left to right from the most severe to the less severe ones, consistently with the typical behavior of short term memory, the number of most severe alarms was always correctly remembered, being these the first in the list, as well as in the 10 and 12 cases.

Table 6. Results of the scalability analysis.

| Nr (K) of Involved NEs | Max Subareas per NE \(^1\) |
|------------------------|-----------------------------|
| 1                      | 8                           |
| 2                      | 8                           |
| 3                      | 5                           |
| 4                      | 4                           |
| 5                      | 3                           |
| 5 < K < 16             | 1                           |

\(^1\) i.e., threshold for number of simultaneous active alarms in the same NE.

In the case of inter-NE simultaneity, two cases must be considered, depending on whether each NE is the source for at most one alarm or more than one. As to the glance-ability of the synthetic component, the former case leads back to the intra-NE simultaneity case, while for the latter it is necessary to test different configurations corresponding to an increasing number of involved NEs (Figure 7b illustrates the cases of two/three/four distinct NEs). Since for the inter-NE case, the notifier visualizes alarm information considering in sequence one NE after the other, it is possible to analyze the synthetic component of each NE area independently of what happens in the other NEs involved. In the case of two NEs (illustrated in Figure 7b) we achieved the same performance results as in the intra-NE case (correctness above 80% up to 8 subareas and harsh discontinuity for the 10 case), and we hence fixed a threshold of 8 subareas (see Table 6). With more than 2 NEs, because of the increasing limitation of the available space estate for each NE in the synthetic component the steep degradation of recall performances occurs before than in the two NEs case, leading to a threshold lowering (Figure 7c,d show the cases of three and four NEs, respectively). In particular, for three NEs, four NEs, and five NEs we found threshold values of five, four, and three, respectively (see Table 6), while, for a number of NEs K greater than five, partitioning the NE area of the synthetic component even only
in two portions raises readability issues since the adjacent areas would be so small as to be distinguishable only in case of highly contrasting colors. The maximum acceptable number of involved NEs is 16, which, as to subarea discernibility, is analogous to the case of two involved NEs with eight simultaneous alarms each. Summarizing, we obtain the constraints in Table 6.

In summary, with reference to Table 1, after the 2nd iteration objective PO2 was considered “almost” achieved (but for the flag issue), and objectives RO2.1 and RO4 were considered achieved while objective RO2.3 (and consequently RO2.3) required additional work. A usability test focused on the validation of the mapping between severity levels and transition types was scheduled at the end of the study, after the generalization of results and the identification of possible extensions aimed at overcoming the limits of Table 6.

3.3. Third Iteration: Final Release of the Proof-of-Concept and Definition of a General Formal

With the substantial success of the cognitive walkthrough session and success of the workshop, the company interest was by and large completed at the end of the 2nd iteration. Hence, beside amending the flag flaw, the 3rd iteration was primarily focused on the research interest and on the achievement of research objective RO3 (Release of a general abstract model for the notification of multiple alerts from heterogeneous sources).

3.3.1. Final Release of the Basic GLANCE Model and of the Proof-of-Concept Mock-Up

To solve the detected flag issue, the detailed component was refined by replacing flags with annotated tabs (codifying severity by color, size, and LOD) not overlapping with the synthetic component, with the additional benefit that, in an interactive tool, tab affordance would allow operators to manually access on-demand information on each individual alarm on the same NE, independently of its severity, as brought to light by the expert-based evaluation. The new visual signals and an example situation are illustrated in Figure 8.

![Figure 8](image_url)

*Figure 8. Redefinition of multiple fault information in the detailed component: (a) The new intra-NE simultaneous alarms signals (color, size, and LOD depends on the severity level) and (b) an example configuration of a hybrid case.*

Given that the usability flaw was related only to the overlapping of color and that the new introduced signal avoids overlapping, objectives RO2.1 and RO2.3 was considered satisfied after an inexpensive expert-based usability test. Furthermore, a Java prototype was implemented according to the model defined so far and objective PO2 was considered satisfied.

3.3.2. GLANCE Generalization

At this stage of the design, the general GLANCE model can be described by results summarized in Figure 3 (single alarm visual coding technique), Table 5 (transition strategy), Section 3.2.1 (multiple alarms overview technique), and Table 6 (simultaneity threshold). Though the initial goal was domain specific, the achieved design exhibits a level of abstraction that stimulated to investigate not only improvements aimed at overcoming the limits single
out by the scalability analysis, but also the applicability of GLANCE in other contexts 
where the information to be notified satisfies the few basic characteristics identified for the 
alarms studied so far: (1) Classification along the dimensions of class and severity, and 
(2) description as tuples of values.

While the threshold on simultaneity proved not to be a limitation for the specific 
optical fiber networks the company had to deal with (given typical topology and number 
of elements of such networks), the scalability analysis suggested that extensions were 
advisable to obtain a broader applicability, both in more general TLC networks and in 
additional domains. In a generalizability perspective, the ultimate objective was the 
capability of notifying alerts/updates from different heterogeneous information channels, each 
potentially generating a high number of notifications. Moving from the specific TLC 
network application domain to general application domains, the limitations dictated by 
the scalability analysis on the number of involved network elements and on the number 
of alarms active in the same network element can be reformulated as limitations on the 
number of different channels and on the update rate of individual channels, with an impact on 
the display efficacy.

A first natural extension is the introduction of hierarchization mechanisms with the 
twofold positive effect of mirroring the customary hierarchical organization of several real 
application domains (e.g., news organized by topics) and providing a mechanism that 
possibly breaks down complex cases into smaller more manageable pieces (e.g., traffic 
information on Italian highways may be dispatched on a regional basis), as well as in 
the primary application domain of the study (e.g., a TLC may be structured in smaller 
subnetworks).

A first redesign step was hence aimed at studying new overviewing mechanisms for si-
multaneous alerts, oriented to information classification and hierarchization, general enough 
and orthogonal, to be selected and possibly combined so as to adapt the display to channels 
to dispatch. The GLANCE extension was based on the following assumptions:

- To conform to previous design choices regarding the synthetic component partitioning,
- To organize the “notification space” into simpler sub-spaces so to have a lower proba-
bility that a single sub-space exceeds the threshold,
- To display the status of one sub-space at the time, and
- To offer mechanisms to switch among sub-spaces.

The two newly introduced orthogonal mechanisms are horizontal scrolling and display 
tabbing, built on top of the structure considered so far, each exhibiting pro and cons, 
supporting respectively the management of homogeneous elements belonging to the 
same category (e.g., traffic alerts from the highway infrastructure) and of notification 
spaces organized in distinct categories (e.g., folders of email). Figure 9 illustrates some 
example scenarios:

- The display in Figure 9a illustrates a possible use of horizontal scrolling in the case of a 
  highway traffic monitor: Regions with incoming news are treated separately (in the 
same way of network elements) and shown in sequence, each presenting simultane-
ous news by customary synthetic component partition (in the same way of alarms 
generating in the same NE). Arrows at the extremes of the synthetic component act 
as affordances for its behavior (anyway, in the default ambient behavior the notifier 
would autonomously and continuously scroll the list, moving from element to ele-
ment). This scenario corresponds to monitoring a set of elements \( S = \{E_i, i = 1, \ldots, n\} \), 
where each \( E_i \) acts as a source of multiple notifications and can hence be viewed as an 
alternative solution for the hybrid situation of Case 3 in Section 3.2.2: The advantage 
is the absence of limit in the number of scrollable elements, but its disadvantage is 
the loss of a global overview and the lack of direct access to individual regions. It 
is worth noting that the notifier allows one to scroll sets of elements (each managed 
as in the basic model discussed for network management), according to two-level 
multichannel scenarios;
• The display in Figure 9b illustrates the use of display tabbing, showing notifications of incoming email, organized in categories (e.g., inbox, spam, work, etc.) represented by tabs. This solution has the advantage of offering a glanceable global overview and guaranteeing an easy direct access to individual categories (tabs) in an interactive tool (again, in the default ambient behavior the notifier would autonomously and continuously move from tab to tab). The disadvantage is the upperbound on the number of categories determined by the limitation on the display width. Another disadvantage can be the possible contiguity of red/orange and green tabs, which would require additional visual cues for tabs with major/critical notifications (e.g., thicker borders);

• The display in Figure 9c illustrates the combined use of the two mechanisms in a two-level multichannel scenario in which individual channels are associated to tabs and are further structured into subcategories handled by the horizontal scrolling of the display.

Figure 9. Handling multiple notifications: (a) An example of horizontal scrolling in the case of highway traffic information organized by regions; the currently displayed region is Puglia, with two incoming notifications; (b) an example of tabbing in the case of incoming mail; the currently displayed category is “Spam”, with one incoming notification; and (c) an example of a two-level structure: The currently displayed category is “News”, with four notifications scattered into a number of subcategories handled by horizontal scrolling; the currently displayed subcategory is “Sport”, with two incoming notifications.

It must be observed that the utilization of such mechanisms allows one to overcome the limitations on the threshold given in Table 6. For example, a TLC network partitionable in subnetworks may be managed either by assigning subnetworks to different tabs or by simply scrolling the subnetworks.

In summary, the two additional mechanisms of horizontal scrolling and tabbing extend GLANCE monitoring/overviewing capability and allow one to manage structured notification spaces. The features of GLANCE can be the basis for customizable multichannel notifiers built and configured onto multisource prioritized notification dispatchers. As the discussion about the example scenarios showed, configuration choices should be on a case-by-case basis, taking into account the characteristics of the monitored notification sources and the degree of complexity of the notification space, in order to select the best combination of overviewing mechanisms.

3.3.3. Evaluation and Learning

As discussed in Section 3.1.1, the design choices related to transitions in Table 5 are based on already validated literature results. In particular, the five severity levels used in GLANCE (cleared, warning, minor, major, and critical) are mapped onto four notification levels (change-blind, make-aware, interrupt, and demand-action) adopted and validated by PTK, a toolkit providing structured support for managing user attention in the develop-
ment of peripheral displays [38], and associated to transition types known to be adequate for making displays work in the correct zone of the attention-awareness graph (inattention, divided attention, and focused attention), as surveyed in [38]. In particular, displays working in the inattention zone use very gradual animation, minimizing motion (e.g., [38,64]), displays working in the divided attention zone employ a number of techniques based on abrupt changes (e.g., [38,65,66]), and displays working in the focused attention zone use significant motion/changes in the interface to draw user’s attention (e.g., [38,67]).

Having based transitions on such validated results, we expected GLANCE to correctly behave as a change-blind/ambient/alerting display based on notification severity, and for this reason the evaluation at the end of the first two iterations was primarily focused on the GLANCE visual coding technique, which, as already observed, is the leading aspect of the study. At the end of the third iteration, we set up a usability test to evaluate the extended model with scrolling and tabbing mechanisms including also the validation of the mapping in Table 5, i.e., the GLANCE ability to make users “consume” notifications in the correct way, with an immediate reaction to severe ones and subliminal long-term comprehension of minor ones.

**Experiment design.** The scenario in Figure 9c, which includes both new mechanisms, was evaluated by a laboratory-based controlled observational study grounded on a dual-task test to simulate what happens in a scenario in which notifications arrive while the operator is attending other tasks. The objective of the study was to measure how GLANCE notification effects can influence performances in a dual-task setting, both on the foreground and the background task. The study was designed to have participants performing a primary text-reading task while, at the same time, monitoring notifications delivered by GLANCE. We wanted to measure, on the one hand, how distracting the notifications were on the text-reading performance (by user observation and a post-experiment test) and, on the other hand, how memorable the notifications were (by a post-experiment test). Notifications to be administered to participants were designed to be credible content of a multichannel notifier. In particular, the sample GLANCE display used for the test included four categories (i.e., four tabs) related to *stock market* (with two subcategories), *incoming email* (with four subcategories), *news* (with five subcategories), and *calendar* (with four subcategories). The notification sequence included information related to 16 different events and a total of 13 animations of display content (Figure 9c is actually an example notification from the tested prototype), with a uniform distribution of notifications and display changes over the four notification levels (change-blind changes, make-aware, interrupt, and demand-action). A pilot test was conducted with internal users to tune the timing of the notifications relatively to the note-taking activities of observers. Based on the results of the pilot test, we decided to have two observers per user, alternating within a single user observation during the sequence of display changes, so that each of them could have enough time for recording details about the user’s reaction to each display change. Observers were trained to judge user’s reaction based on users’ eye movements and face expression in response to a notification and were asked to collect structured quantitative data to be organized as an observation matrix: $displayChange \times user \rightarrow rating$, where rating was to be assigned by the observers according to a four values scale depending on the attention zone corresponding to the user reaction ($0 =$ inattention, $1 =$ divided attention, $2 =$ focused attention, and $3 =$ focused attention plus action). The observation matrix was to be compared with the “expected matrix” after the experiment to measure the distance between actual users’ behavior and expected behavior based on the notification level.

**Experiment sessions.** A total of 15 participants (six female), aged 22–51 (AVG = 33.6; $M = 30$), took part in the study on a voluntary basis, recruited among students and personnel of the university campus (five participants were Master students, two PhD students, two laboratory technicians, two were from administrative personnel, and four were employees not using computers for their jobs). None of the participants had previous experiences with the tested notifier, but seven of them had been involved in the scalability test of the previous prototype. Participants were invited individually to the lab and were
given a brief introduction to the experiment. A facilitator gave a description of GLANCE features and explained to users what they were requested to do, what they would be observed for, and how data collected from the study would be used. Then each user was asked to read a one-page text (foreground task) while simultaneously monitoring the asynchronous notifications delivered by the GLANCE display (background task). In the case of critical notifications (corresponding to the demand-action level) they were requested to click on the display, while in all other cases they were not asked to do anything in particular. Participants were not instructed on how to prioritize the two tasks nor were they preliminarily informed about the speed with which the notifier would update (actually, notifications arrived in an asynchronous manner). At the end of the session, participants were asked to recall details of both the read text and information notified by the GLANCE display by answering a questionnaire with 12 questions about event information delivered by GLANCE during the test. Furthermore, in a post-experiment debriefing, participants were asked to subjectively evaluate GLANCE as to intuitiveness, glanceability, intrusiveness, and usefulness of the display.

**Results.** At the end of the test sessions we analyzed the data collected by the observation matrix and questionnaires. As to the former, gathered data were compared with the expected values and discrepancies were analyzed and summarized (Table 7). In particular, we observed that anomalies in the observed attention zone involved a distance of only one level up (e.g., a “divided attention” was observed when an “inattention” was expected) or one level down (e.g., a “divided attention” was observed when a “focused attention” was expected). Only in one case was there a recorded distance of two levels down, but the associated observer’s note indicated that it occurred when the user was distracted from the foreground task. Table 7 summarizes the results, differentiating between the two cases of “level up” (the display change grabbed more user attention than due) and “level down” (the display change grabbed less user attention than due).

| Notification Level   | Rate of Level Up | Rate of Level Down |
|----------------------|------------------|--------------------|
| Inattention          | 13%              | 0%                 |
| Divided attention    | 8%               | 11%                |
| Focused attention    | 4%               | 8%                 |
| Demand action        | 0%               | 12% (6%)           |

We notice that overall the display behaved as expected. It is worth nothing that, if we consider the total of inattention and divided attention display changes (90 changes), we detected unwanted interruptions only in 4% of cases. On the other hand, if we consider the total of focused attention and demand action display changes (105 changes), we did not get the desired interruptions in only 3.8% of cases. As to the rate of level down in the demand action notification level, it has to be noticed that more than 50% of cases occurred in the first appearance of a demand action display change (these users did promote the background task to foreground task but, without instructions, they forgot to click on the display as requested by the test). If we ignore the first demand action display change for all users and we limit the calculation to the remaining display changes, the rate drops from 12% to 6%. We also notice that, while in the administered test the requested action did not produce any particular benefit to the users, in real contexts it would be associated to some domain-specific meaningful effect that would push users to perform it. Actually, in the post-experiment debriefing, some users reported that they perceived the requested action as a “chore” and they would have liked, e.g., to access the source of the critical news by clicking.
As to the users’ post-test, the result of the questionnaires showed that the notifier did not affect at all the completion of the primary task (with a correctness rate of 100% on the questions about the proposed text). As to the secondary task, Table 8 summarizes the correctness rate (CR) of users’ answers, rated according to four levels: Excellent (CR > 90%), good (70% < CR ≤ 90%), sufficient (50% < CR ≤ 70%), and poor (CR < 50%).

Table 8. Results of the secondary task of the dual tasks test.

| Id | Question                                                                 | CR  |
|----|--------------------------------------------------------------------------|-----|
| 1  | How many categories (tabs) were monitored by the notifier?              | excellent |
| 2  | Which category in the following list was not monitored?                 | excellent |
| 3  | How many critical alarms do you remember?                              | excellent |
| 4  | For which categories do you remember critical alarms?                  | good  |
| 5  | Which among the following stock market indices was involved in a notification? | poor  |
| 6  | How many email notifications do you remember?                           | sufficient |
| 7  | Which was the sender’s name of a notification in the group “Personal”?  | sufficient |
| 8  | Do you remember notifications about a sport event and if yes its score? | good  |
| 9  | Which calendar events do you remember?                                 | excellent |
| 10 | Do you remember political news notifications and if yes which?          | excellent |
| 11 | Do you remember what the weather notification was about?               | good  |
| 12 | Which was its background color?                                        | sufficient |

Overall, the correctness rate of the secondary task was above 80% in 66% of the questions (1–4, 8–11), and around 60% in the remaining 33% (5–7, 12). It is worth noticing that user performances reflect somehow the results obtained by the observational study, which showed that some ambient notifications fell in the inattention zone. In particular:

- There was no detection of statistically relevant differences based on age or computer experience;
- Notwithstanding that 8% of alerting notifications fell in the ambient zone, critical information was always correctly recalled;
- Most significant mistakes are related to the recall of some specific details of a low-to-medium severity notification (like, e.g., the name of a stock index in Question 5, the sender of an email message in Question 7, the background color of a panel in Question 12);
- Question 12 was asked just to check how many different information users were able to digest about a notification but remembering a notification color is not the objective of the system.

As to the lower memorability of lower level notifications compared to higher level notifications, it has to be observed that the test articulation (text reading and monitoring first, and memory verification after) did not emulate exactly what would happen in a real scenario in which users would react somehow when the notification arrives in case of high severity notifications (which are proved to be all very well recognized and recalled), while they would perceive in a subliminal way less severe notifications that will be cyclically displayed by the notifier, thus not necessarily requiring long term memorization. Therefore, lower users’ performances on lower severity notifications of questions 5 and 7 do not correspond to actual bad performances in a real setting and are not to be considered worrying.

As to the post-experiment debriefing, participants were asked to provide scores on a 5-value scale (from 0 to 4, the higher the better) in relation to their subjective judgement about (1) the overall intuitiveness of the visual coding technique, (2) capability of providing a glanceable overview of the notification space, (3) perceived lack of intrusiveness of dynamic stimuli, and (4) usefulness of the display in routine working activity. As reported in Table 9, in this case we obtained different feedback based on the role of participants, with students assigning a slightly better rating to intuitiveness and glanceability of the display,
and campus personnel assigning higher scores to lack of intrusiveness and usefulness of the display.

Table 9. Subjective GLANCE evaluation.

| GLANCE Property         | AVG of Students | AVG of Personnel | Overall AVG |
|-------------------------|-----------------|------------------|-------------|
| Intuitiveness           | 4.29            | 4                | 4.13        |
| Glanceability           | 4.29            | 4                | 4.13        |
| Lack of intrusiveness   | 3.86            | 4.12             | 4           |
| Perceived usefulness    | 3.57            | 4.37             | 4           |

As to intuitiveness and glanceability, we believe that the slightly higher scores of students may be due to the greater familiarity of (digital native) younger people to visual tools. More interesting are the results relative to the possible integrability of GLANCE with routine tasks, where the frequency and obligation of attending primary tasks with expected productivity probably make a difference, particularly in the perceived usefulness of a product capable of minimizing unwanted interruptions while guaranteeing information awareness. This result is in line with the expert-based evaluation and with the discussions with Technolabs personnel during the workshop, and in line with the overall primary objective of GLANCE.

Based on all these considerations, objective RO3 was considered satisfied and the project was considered completed.

4. Results of the 2nd Macro-Phase: Application and Refinement of GLANCE

In this section we briefly survey the results of the second design macro-phase, which included two traditional UCD-based projects which allowed us to evaluate the flexibility of GLANCE to adapt to different contexts. The first project studied the application of GLANCE to a personal customizable multichannel notification tool and led to an extended version of GLANCE, denoted GLANCE+, including additional interactive features coherent with the Shneiderman’s mantra (overview first, zoom and filter, details on demand) [68]. The second project adapted GLANCE and GLANCE+ features to a service-oriented fault monitor for digital terrestrial television broadcasting networks.

4.1. Application to a Personal Customizable Notifier

As observed in Section 3.2.2, the GLANCE monitoring/overviewing capability, extended by the introduction of scrolling and tabbing mechanisms, allows one to view the features of GLANCE as the basis for customizable multichannel notifiers. We hence investigated the exploitation of achieved results for designing a personalizable multichannel notification display. In particular, we considered a notification space composed of a number of distinct categories (e.g., news, email, weather, etc.) as in the scenario depicted in Figure 9c (categories are associated to tabs and further structured into subcategories handled by scrolling). An additional analysis led to a revision of some aspects of the category management, aimed at overcoming two different efficacy problems. The first one is related to the abstractness of the display with respect to category management, which, though formally correct, may cause “similarity interference” (a problem discussed also in [69]): When formatting visual displays for dynamically updating environments, the design has to make information highly distinctive across the displayed items. In our case, while the currently visualized category is easily recognizable, no information other than number and severity of incoming notifications is graspable at a glance for the other categories. The second problem is related to the upper bound on the number of manageable tabs, which translates into an upper bound on the number of categories that users may decide to monitor, hence becoming a sever limitation to the broader applicability of the notifier.

A unified solution for the two problems was found by rethinking the category area according to the basic ideas in Figure 10: A category bar behaves as a scrollable list of categories visually represented by icons (Figure 10a), with the multifold effect of removing
the upper bound on the number of categories, solving the similarity interference problem, and exploiting perceptual sensory immediacy. Furthermore, as illustrated in Figure 10b, a category is associated to a status sensitive icon, i.e., to a family of icons getting visually richer as the number of simultaneous alarms increases, thus providing a rough indication of the number N of simultaneous incoming notifications within a category (the exact indication of N is provided by tagging the icon with N). In other words, we now combine in the category bar most of the information previously split in the tabs and in the synthetic components. Clearly, the icon style is just a matter of aesthetic and the one in Figure 10 is to be viewed just as an example of the new approach based on status sensitive icons (actual tools based on this model might also offer a choice of multiple themes). Notice also that status sensitive icons allow now a first LOD exclusively based on the category bar showing a glanceable overview on the number of incoming notifications (simple additional visual hints might also indicate which categories include critical notifications, if any) and the small arrow below the icons allows one to access a second LOD including the notification details as in the previous version of GLANCE.

Figure 10. The new category bar: (a) Its organization, and (b) an example of status sensitive icon.

Now, since in an application like this it is reasonable to offer both a user-initiated and a system-initiated interaction, we need a mechanism to automatically cycle over categories with incoming notifications. To this aim, we conceived, and tested with users, four different solutions (shown in Figure 11), differing in:

- Visual hints used to indicate the current category (tab metaphor vs. lens metaphor),
- Relative movements of categories and overlapping tabs/lens,
- Scrolling direction of individual notifications within the detailed component of a single category (vertical vs. horizontal).

**Experiment design.** The new version of the GLANCE model (denoted GLANCE+) was evaluated by a laboratory-based controlled observational study grounded on a dual-task test similar to the one described in Section 3.3.3. The primary objectives of the study were (1) to evaluate the novelties introduced in GLANCE+ (and hence the glanceability of the new category bar), and (2) to select the solution with the best objective and subjective performances. In this case we did not repeat the evaluation of the mapping between notification levels and transition types, which was considered validated by the results of the previous test. The study was designed to have participants performing a primary text-reading task while, at the same time, monitoring notifications delivered by GLANCE+.

In particular, we wanted to measure how distracting the notifications were on the text-reading performance (by user observation and a post-experiment test), how memorable the notifications were (by a post-experiment test), how intuitive the category bar features were (by user observation and a post-experiment test), and how satisfied users were with the different proposed solutions (by a post-experiment test). Notifications to be administered to participants were designed to be credible content of a multichannel notifier including up to six different channels. To evaluate the efficacy and glanceability of the different alternatives, different rounds of dual-tasks tests were designed, including (a) texts different in content but analogous in terms of complexity and (b) notification sequences different in content but analogous in terms of kinds of notifications and display changes (to compare the different solutions under the same conditions in terms of user familiarity with the content). Notice that for the same solution, two scenarios were designed, differing in the
number of channels included: One scenario with a number of channels not requiring the scrolling of the horizontal bar and another scenario with a number of channels exceeding the number of visible channels in the category bar and hence requiring horizontal scrolling.

![Figure 11](image-url)

**Figure 11.** Four alternatives for the display: (a) A first type of tab-based category bar with horizontal scrolling of individual notifications; (b) a second type of tab-based category bar with horizontal scrolling of individual notifications; (c) lens-based category bar with horizontal scrolling of individual notifications; (d) lens-based category bar with vertical scrolling of individual notifications.

**Experiment sessions.** A total of 10 participants (five female) aged 18–50 (AVG = 32.9; M = 32.5), heterogeneous with respect to technical skill, recruited among students and personnel of the university campus, participated in the study on a voluntary basis (one participant was an undergraduate student, two participants were Master students, two were PhD students, two were technicians, and three were employees not using computers for their jobs). None of the recruited users participated in previous evaluation sessions of GLANCE. Participants were invited individually to the lab and were given a brief introduction to the experiment. A facilitator gave a description of GLANCE features and explained to the users what they were requested to do, what they would be observed for, and how data collected from the study would be used. Then each user underwent different dual-task rounds corresponding to different category bar design alternatives, with the order of the design alternatives assigned randomly to each user, again to compare the different solutions under the same conditions in terms of user familiarity with the content and display. As in the previous usability test, in this case participants were also not instructed on how to prioritize the two parallel tasks nor were they preliminarily informed about the speed with which the notifier would update. At the end of each round, participants were asked to answer a questionnaire including three sections: Questions about details of the read text, questions about details of the information notified by the GLANCE+ display, and questions about features of the category bar. Furthermore, at the end of the entire session, users were asked to subjectively evaluate which alternative they judged as the most intuitive.

**Results.** As to the recall of text details and notification details we expected and indeed achieved a confirmation of results obtained for the previous prototype. It is worth observing that, for both the foreground and background task, in the first administered round the users’ recall performances results were slightly lower than the recall performances of the participants in the previous test (the correctness rate was 91% for the read text and 52% for lower-to-medium severity notifications), while equalizing them in the successive rounds. As reported also by the observers, in this case users required more time to become familiar with display features and behavior but, as the results of successive rounds shows, the system was quickly learned. Once the users had learned the display, their recall performances improved to match the glanceability of the previous version of the display (actually, the administration of several rounds to each user allowed us to get an indication on the learnability of the display). We also observed that in this case information about critical notifications was always correctly recalled, as was in the first rounds. As in the previous test, as to recall performances we did not detect statistical differences based on age or on computer experiences.

As to the specific evaluation of the category bar, one of the novelties to be validated was the use of a combination of status sensitive icon and numeric label to indicate the number of
incoming notifications for a given category: 60% of the users understood that the category icon changed based on the number of notifications, 70% of the users understood the meaning of the numeric label, and all users recognized at least one of the two visual hints as indication of the number of incoming notifications for the category. This confirms the appropriateness of relying on a combination of icon and numeric label.

The other aspect to be validated was the visual hint used to indicate the current category (tab metaphor vs. lens metaphor). As to the selection of the best solution, this was based on both objective criteria, with respect to users’ ability to digest a set of categories organized as a scrollable list, and subjective criteria (i.e., users’ preferences). From the results of the third section of the questionnaire, it emerged that most users did not understand that in the tabbed solutions of Figure 11a,b the category list was a scrollable list (probably because of the familiarity with unscrollable tabbed displays), while the same problem was neglectable in the third and fourth options of Figure 11 (differing only in the scrolling direction of individual notifications within the detailed component of a single category), which had also better performances in regards to the recall of notifications with medium-to-high severity and to the recall of categories overview. Among these two candidates, the organization subjectively preferred by users was the one in Figure 11d, characterized by:

- A lens in a fixed central position of the category bar,
- The category list scrolling from right to left, and
- The notification descriptions scrolling from bottom to top.

Again, as already observed about the icon style, the lens is just a matter of aesthetics and may be replaced also by different (and possibly more abstract) visual solutions coherent with the model (i.e., in a fixed central position of the category bar over a scrollable category list).

The final design of the GLANCE+ model is illustrated by the sample display in Figure 12a. Considering the expectations of users of a personal notifier, the display has been enriched by a scrollbar on the right side to directly interact with the notification list and an interactive control palette on the left side with icons for (a) pausing/restarting the automatic scrolling, (b) going directly to a notification source, and (c) deleting a notification from the list. The list is also visualizable in an on-demand panel, to provide quick direct access to individual selected items (see Figure 12b), thus offering an additional glanceable overviewing capability. A GLANCE+ based Java application has also been implemented, equipped with a back-office component allowing users to customize the notification aggregator.

![sample display](image)

**Figure 12.** The final design of the GLANCE+ model, with (a) the notification component, and (b) the notification list visualized in the on-demand panel.
4.2. Application of GLANCE and GLANCE+ to a Service-Oriented Fault Monitor for Digital Terrestrial Television Broadcasting Networks

The second application project was carried out as a customary UCD-based project within the framework of co-operation between Univaq and Broadcast Solutions S.R.L., an Italian company offering system integration capabilities for broadcasting and media industries. We observe that an exhaustive description of this project goes well beyond the scope of this paper since in this case the primary company objective was to conceive a novel perspective for the fault management approach. We report on this project here, sketching its main features and results, because it was an opportunity for testing the adaptation of GLANCE and GLANCE+ features to a real context.

In particular, the company’s goal was to study a novel approach for fault management in digital terrestrial television broadcasting networks, shifting from a traditional device-based perspective, with scarce support for detecting how the final service is affected by the alarmed device, to a service-based perspective centered around faults as perceived by the final user (e.g., absence of the audio component in a given channel) without losing the possibility of clearly identifying the corrupted device(s). This issue has implications on both the “internal engine” managing the service/device correlation and the “dispatching display” alerting operators about alarms and offering efficient navigation in alarm data. As to the “dispatching display”, after a series of brainstorming meetings with domain experts and final users, it was agreed to study a GLANCE-based fault notification component offering the operator a glanceable overview of information about incoming alarms, grounded on the final users’ service-based perspective (e.g., “absence of the video component in RaiUno”), while guaranteeing a smooth and efficient switch to the device-based perspective (e.g., “lack of video signal on SDI input interface”). Though similar to the original project with Technolabs, the new application has to take into account multiple requirements:

- The fault management package has to provide a service-centric overview to offer the operator a precise view of what the consumer is receiving;
- The fault management package has to provide a service/device correlation to make the operator able to intervene for fixing the problem;
- The alarm notification has to grasp the correct quantity of operator’s attention, depending on alarm severity, and alarm data useful for fast intervention planning have to be easily identifiable.

It must be observed that in digital terrestrial television broadcasting networks, transmission malfunctioning may depend on a number of related devices: Fault on one device may provoke faults on other related devices, making it appropriate to provide a clear correlation amongst all faults affecting the same service. Furthermore, processes carried out by a device are somehow “children” of previous processes and “parents” of processes carried out by successive devices in the chain. It is therefore appropriate to provide a hierarchical visualization of the network status through an “alarm tree” that allows operators to easily understand the correlation service/alarm and to quickly identify a malfunctioning final cause (Figure 13 shows a portion of a simple alarm tree on a given service).

Since in this case the notifier has to deal with one category only, the basic idea was to visualize a glanceable low-LOD overview of alarmed services (basically adhering to the horizontal scrolling case of GLANCE (Figure 9a)), associated to higher-LOD views allowing for the visualization and exploration of insight data of the service alarms accessible on demand as in GLANCE+. In particular, as illustrated in Figure 14a, the upper bar allows scrolling alerted services either in a user-initiated modality or by automatic cycling, the middle area provides an alarmed service overview along with summary information on device faults of its service components, while the lower bar (initially conceived for generic “user data”) is used to provide a service-centric network overview. The on-demand panel introduced in GLANCE+ in this case provides details on service alarms visualized either as a list (Figure 14b) or as a browsable alarm tree (Figure 14c).
The alarm notification has to grasp the correct quantity of operator’s attention, and it must be observed that in digital terrestrial television broadcasting networks, transmission malfunctioning may depend on a number of related devices. The GLANCE-inspired “Service in alarm” fault notification display: (a) A sample glanceable low-LOD overview of an alarmed service providing summary data (number and severity) on its components and their associated alarms on devices, and a service-centric network overview in term of summary of total services in alarm. Operators can access views with higher LOD by clicking on the small arrow at the bottom of the service overview, like in GLANCE+. In this case, this action discloses a tabbed on-demand panel providing details on service alarms either as a list (Figure 14b), or as a browsable alarm tree as in (c) (clicks on tree items provide an additional LOD on devices in a balloon fashion).

The notifier prototype was implemented as a Java application. Usability evaluations included expert-based methods (cognitive walkthrough) and user-based methods (observational evaluation and verbal protocols [70]). Given that the display was grounded on results achieved and validated by the GLANCE and GLANCE+ projects, in this case the evaluation was more centered on specific application domain goals, efficacy of the novel service-based approach, and alarm data navigability, and for this reason is only sketched here since a more extensive discussion would require to address specific service-oriented fault management issues. In particular, expert-based evaluation pertained general issues about information presentation, information comprehension, and interface interaction, and singled out issues to be investigated by user-based evaluation. The user-based evaluation sessions were conducted in-situ with five experienced PC users, based on a classical user-based schema. Users were recruited among company technical personnel and were potential actual end users of the prototype (“operators”), all male, aged between 25 and 40, experts at the application domain. The experimental evaluation was conducted in five distinct sessions (one for each user, each with a duration of about two and a half hours) in the real users’ working environment: Users were asked to perform the test during their ordinary job activities (including, e.g., coffee breaks, attendance to other tasks, discussions with colleagues) to test the efficacy of the notifier in the real setting. As discussed...
later on, the notifier was very appreciative and proved to be effective with respect to its operational goals.

5. Discussion

The work presented in this paper can be discussed according to different perspectives: The taxonomic characterization with respect to existing frameworks, the methodological approach, and the industrial acceptance.

As to the taxonomic point of view, GLANCE and its variants can be characterized according to two frameworks consolidated in literature. The study in [71] proposes a taxonomy of ambient and alerting systems based on four design dimensions denoted information capacity, notification level, representation fidelity, and aesthetic emphasis; according to such taxonomy, GLANCE can be defined as an “Information Monitor Display” and is located in a peripheral area of the desktop, can visualize information from a high number of discrete information sources, allows several notification levels, and is not characterized by high aesthetic emphasis. As to the IRC framework [36], the assessment of GLANCE, based on the results of the user-based evaluation, emphasizes its ability to adapt its behavior to the alarm severity:

- For low severity alarms (low interruption) display updates grant longer term awareness (high comprehension) of notification items without provoking immediate reaction (low reaction). In this case the system is mapped to the 001 IRC triple behaving like an ambient display (located in the lower part of the awareness-attention graph of Figure 5) functioning at “operational level” according to the activity theory discussed in [34]: Experiments showed that non critical information was remembered with acceptable precision and user observation showed that its arrival did not promote the notifier to the primary task of the users, as expected (as shown by results in Table 7);

- For higher severity alarms, display updates require higher amount of attention, up to requiring urgency awareness (high comprehension), interruption of all ongoing tasks (high interruption), and quick intervention (high reaction). The system is therefore mapped to the 111 IRC triple behaving like an alerting display (located in the higher part of the awareness-attention graph of Figure 5) functioning at “action level” according to the activity theory discussed in [34]: Experiments showed that critical information was always identified, recognized, and remembered, and users observation showed that the arrival of major and critical notifications made the notifier promote itself to the primary task of the users, as expected (as shown by results in Table 7). The efficacy of high severity notifications was proved also by the GLANCE workshop at Technolabs, as discussed later on.

With respect to methodological aspects, we want to underline the positive experience of the contamination of HCI methods and techniques with the tandem model idea borrowed from the Action Research realm and utilized as a general organization guidance of the work in the first macro-phase. In particular, the clear identification of research objectives on one side and problem objectives on the other side, as well as the explicit acknowledgment of research interests on one side and problem interests on the other side, proved to be greatly beneficial for the cooperation between the project partners and for a rigorous and structured project management taking care of all involved stakeholders. The positive effects are also testified by industrial satisfaction.

As to the industrial acceptance point of view, related to the initial objective of overcoming limitations of notifier components of commercial network monitoring tools, the main challenge of the project was the shift of paradigm to peripheral displays in applications addressed to traditional and possibly conservative working contexts. Actually, both in the project in cooperation with Technolabs and in the application project in cooperation with Broadcast Solutions, the proposed conceptual solutions and the achieved results received great appreciation, well beyond expectations. As to the first project, the GLANCE workshop organized by Technolabs people was a unique briefing and testing opportunity. To our surprise company managers showed great interest in not only in the problem-solving side
of the project but also in the research side. Furthermore, company people reactions to the proof-of-concept demo provided additional evaluation elements, with a clear confirmation of the notifier suitability at provoking reactions correctly depending on the alarm severity. As to the application project, results of the evaluation sessions at Broadcast Solutions were very positive with all users having paid attention to high severity alarms and a significant role was played by auditory stimuli since operators had to leave their desktops frequently to attend other ongoing working tasks; all operators easily used the alarms tree, discovered HW problems related to alarms, and detected HW/alarm correlations, in particular from a “service in alarm view”; all users easily interacted with the interface both at service-oriented level and at device-oriented level; as to subjective judgment, all users considered the interface usable, considered the information organization “well arranged”, and judged the display intuitive as to the visualization of alarm data at the glance on both levels; minor concern was expressed only for background color and severity indicators of service components, a flaw easily surmountable by indicating only the total number of alarms associated to a service component. Overall, the proposed service-oriented GLANCE-based fault management notification was hence considered successful.

It has to be observed that in both cases the objective was the definition and evaluation of proof-of-concept prototypes and that the implementation of operative releases of the notifiers was beyond the scope of the projects and left to the company.

6. Conclusions

In this paper we presented the design evolution of GLANCE, a model for severity-based glanceable notifiers based on a peripheral display approach, to support users in the awareness-interruption trade-off, particularly in the network management application domain or, more generally, in application domains where (a) notification severity plays a crucial role and (b) users need a global overview of incoming notifications. More precisely, the study of customary general approaches to notification management shows that they may be inefficient in the specific application domain of network management systems since they deal with streams of individual notifications displayed one at a time according to a unified format and inserted in lists ordered by the time of arrival without any overview capability. To overcome such limitations, we proposed the different perspective of a notifier that retains the information expressiveness of the one-notification-at-the-time perspective while guiding it by severity-based coding strategies and enriching it with overview capabilities, to guarantee possibly subliminal long-term local and global content comprehension and prompt reaction when the interruption from the foreground task is dictated by fault severity.

The design evolution followed a path composed by two macro-phases, respectively focused on model definition and model application. Though the initial goal was domain specific, the achieved design exhibits a level of abstraction that stimulated to investigate its applicability in other contexts, provided that information to be notified satisfies few basic characteristics, both in terms of single notification data and structure of the notification space.

Future work will be focused on improving the fit with mobile devices and on more extensively studying the promising case of personal notification aggregator. Nowadays the availability of virtually infinite information sources as well as the pervasiveness of mobile devices makes information overload a severe multifaceted problem requiring solutions that are able to help users cope with a number of dimensions of a different nature: Information growth and diversity, human-to-information interaction, interferences with ordinary working, and private life activities. As pointed out in [72], “information overload is as much a problem of information diversity, or clutter, as of its quantity”. The paper reported the results of a test conducted with workers in more than 1000 large organizations and when asked which they considered the worse between the quantity of information they have to deal with or its diversity, respondents identified the diversity as the major cause of stress. Though the survey dealt with the general case of mix of digital and paper
information, we find it reasonable that lessons learned from it can also be applicable to the case of digital information only. It is therefore appropriate studying tools in helping users deal more efficiently with continuous and diverse incoming information.

Author Contributions: Conceptualization, L.T., S.D.P., D.A., A.B., and A.C.; methodology, L.T.; software, D.A., A.C., and A.B.; validation, L.T.; formal analysis, L.T., D.A., and A.C.; investigation, D.A., A.B., A.C., S.D.P., and L.T.; resources, D.A.; data curation, L.T.; writing—original draft preparation, L.T., D.A., A.B., A.C., and S.D.P.; writing—review and editing, L.T., D.A., and S.D.P.; visualization, L.T., S.D.P.; supervision, L.T.; project administration, L.T.; funding acquisition, L.T. All authors have read and agreed to the published version of the manuscript.

Funding: Work discussed in Section 4.2 was partially supported by Broadcast Solutions S.R.L. within the framework of an agreement with the University of L’Aquila for students’ stages.

Institutional Review Board Statement: All the experiments mentioned in the paper are customary usability evaluation with no risk, conducted with voluntary healthy adult people, capable of giving informed consent and that could withdraw at any time, within the framework of official Master thesis works of the School of “Ingegneria Informatica e Automatica” of the University of L’Aquila (Italy) approved by the institutional School Review Board [56–59].

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available [56–59].

Acknowledgments: We thank Doriana Guetti and Umberto Innocente from Technolabs S.p.A., Italy, for their cooperation in the work described in Sections 3.1 and 3.2, Luca Sciotti from Broadcast Solutions S.R.L. for his cooperation in the work described in Section 4.2, the Broadcast Solutions S.R.L. technical staff who participated in the usability testing, and Tania Di Mascio from the University of L’Aquila for the support in the usability testing in Section 4.2. We thank the anonymous reviewers for comments and observations that stimulated the revision of the paper. Furthermore, we note that all authors of the paper were with University of L’Aquila while working in the GLANCE research.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Clemm, A. Network Management Fundamentals; Cisco Press: Indianapolis, IN, USA, 2006.
2. 10 Best Network Monitoring Tools & Software of 2020. Available online: https://www.comparitech.com/net-admin/network-monitoring-tools/ (accessed on 29 September 2020).
3. 10 Best Network Monitoring Tools & Software of 2020. Available online: https://www.pcwdld.com/best-network-monitoring-tools-and-software (accessed on 29 September 2020).
4. The Best Network Monitoring Software for 2020. Available online: https://www.pcmag.com/picks/the-best-network-monitoring-software (accessed on 29 September 2020).
5. Streefkerk, J.W.; McCrickard, D.S.; van Esch-Bussemakers, M.P.; Neerincx, M. Balancing awareness and interruption in mobile patrol using context-aware notification. Int. J. Mob. Hum. Comput. Interact. 2012, 4, 1–27. [CrossRef]
6. McCrickard, D.S.; Chewar, C.M. Attuning Notification Design to User Goals and Attention Costs. Commun. ACM 2003, 46, 67–72. [CrossRef]
7. Bailey, B.P.; Konstan, J.A. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. Comput. Hum. Behav. 2006, 22, 685–708. [CrossRef]
8. Adamczyk, P.D.; Bailey, B.P. If Not Now, When?: The Effects of Interruption at Different Moments Within Task Execution. In Proceedings of the CHI 2004, Vienna, Austria, 24–29 April 2004; pp. 271–278.
9. McFarlane, D. Coordinating the interruption of people in human-computer interaction. In Human Computer Interaction INTERACT’99; Sasse, M.A., Johnson, C., Eds.; IOS Press: Amsterdam, The Netherlands, 1999; pp. 295–303.
10. Horvitz, E.; Apacible, J.; Subramani, M. Balancing awareness and interruption: Investigation of notification deferral policies. In User Modeling 2005: Proceedings of the 10th International Conference (UM 2005); Ardissono, L., Brna, P., Mitrovic, A., Eds.; Springer: Berlin, Germany, 2005; pp. 433–437.
11. Bailey, B.P.; Iqbal, S.T. Understanding Changes in Mental Workload during Execution of Goal-Directed Tasks and Its Application for Interruption Management. ACM Trans. Comput. Hum. Interact. 2008, 14, 21. [CrossRef]
12. Robotham, D.; Satkunanathan, S.; Reynolds, J.; Stahl, D.; Wykes, T. Using digital notifications to improve attendance in clinic: Systematic review and meta-analysis. BMJ Open 2016, 6, e012116. [CrossRef] [PubMed]
13. Pielot, M.; Church, K.; de Oliveira, R. An in-situ study of mobile phone notifications. In Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services—MobileHCI ’14, Toronto, ON, Canada, 23–24 September 2014; ACM Press: New York, NY, USA, 2014; pp. 233–242.

14. Shirazi, A.S.; Henze, N.; Dingler, T.; Pielot, M.; Weber, D.; Schmidt, A. Large-scale assessment of mobile notifications. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’14), Toronto, ON, Canada, 26 April–1 May 2014; ACM Press: New York, NY, USA, 2014; pp. 3055–3064.

15. Iqbal, S.T.; Horvitz, E. Notifications and awareness: A Field Study of Alert Usage and Preferences. In Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work—CSCW ’10, Savannah, GE, USA, 6–10 February 2010; ACM Press: New York, NY, USA, 2010; pp. 27–30.

16. Priyanka, P.; Sivapriya, T.; Laheen, S.D.; Lahari, V.L. SMS Notifications using Raspberry pi. Int. J. Adv. Res. Sci. Eng. Technol. 2019, 6, 8515–8519.

17. Riadh, M.H. Notification System to Students using an Android Application. Int. J. Comput. Appl. 2016, 140, 22–27.

18. Weber, D.; Shirazi, A.S.; Henze, N. Towards Smart Notifications using Research in the Large. In Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct—MobileHCI’15, Copenhagen, Denmark, 24–27 August 2015; ACM Press: New York, NY, USA, 2015; pp. 1117–1122.

19. Kubitz, T.; Schmidt, A. Towards a Toolkit for the Rapid Creation of Smart Environments. In Proceedings of the 5th International Symposium on End-User Development (IS-EUD ’15), Madrid, Spain, 26–29 May 2015; Springer: Berlin, Germany, 2015; pp. 230–235.

20. Kubitz, T.; Voit, A.; Weber, D.; Schmidt, A. An IoT infrastructure for ubiquitous notifications in intelligent living environments. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp ‘16), Heidelberg, Germany, 12–16 September 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 1536–1541.

21. Cutrell, E.; Czerwinski, M.; Horvitz, E. Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance. In Human Computer Interaction INTERACT2001; IOS Press: Amsterdam, The Netherlands, 2001.

22. Czerwinski, M.; Horvitz, E.; Willhite, S. A diary study of task switching and interruptions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’04), Vienna, Austria, 24–29 April 2004; ACM Press: New York, NY, USA, 2004; pp. 175–182.

23. Iqbal, S.T.; Bailey, B.P. Effects of intelligent notification management on users and their tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’08), Florence, Italy, 5–10 April 2008; ACM Press: New York, NY, USA, 2008; pp. 93–102.

24. Cramer, H.; Goddijn, J.; Wielinga, B.; Evers, V. Effects of (in) accurate empathy and situational valence on attitudes towards robots. In Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, Osaka, Japan, 2–5 March 2010; ACM Press: New York, NY, USA, 2010; pp. 141–142.

25. De Vries, R.A.J.; Lohse, M.; Winterboer, A.; Groen, F.C.A.; Evers, V. Combining social strategies and workload: A new design to reduce the negative effects of task interruptions. In CHI ’13 Extended Abstracts on Human Factors in Computing Systems (CHI EA’13); ACM Press: New York, NY, USA, 2013; pp. 175–180.

26. Czerwinski, M.; Cutrell, E.; Horvitz, E. Instant messaging and interruption: Influence of task type on performance. In Proceedings of the OZCHI 2000, Sydney, Australia, 4–8 December 2000.

27. Horvitz, E.; Apacible, J.; Subramani, M. Balancing awareness and interruption: Investigation of notification deferral policies. In Proceedings of the UM ’05, Edinburgh, Scotland, 24–29 July 2005; Springer: Berlin, Germany, 2005.

28. Fischer, J.E. Studying and tackling temporal challenges in mobile HCI. In CHI ’10 Extended Abstracts on Human Factors in Computing Systems (CHI EA ’10); Association for Computing Machinery: New York, NY, USA, 2010; pp. 2927–2930.

29. Abad, Z.S.H.; Marras, O.; Schneider, K.; Barker, K.; Bauer, M. Task Interruption in Software Development Projects: What Makes some Interruptions More Disruptive than Others? In Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018 (EASE’18), Christchurch, New Zealand, 27–29 June 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 122–132.

30. Mankoff, J.; Dey, A. From conception to design: A practical guide to designing Ambient displays. In Public and Situated Displays; O’Hara, K., Perry, M., Churchill, E., Russell, D., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; pp. 210–230.

31. Tabuenca, B.; Wu, L.; Tovar, E. The PRISMA: A Visual Feedback Display for Learning Scenarios. Proceedings 2019, 31, 81. [CrossRef]

32. Colley, A.; Raudanjoki, O.; Mikkonen, K.; Häkkilä, J. Plant shadow morphing as a peripheral display. In Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia (MUM ’19), Pisa, Italy, 27–29 November 2019; ACM: New York, NY, USA, 2019; pp. 1–5.

33. Snyder, J.; Matthews, M.; Chien, J.; Chang, P.F.; Sun, E.; Abdullah, S.; Gay, G. MoodLight: Exploring Personal and Social Implications of Ambient Display of Biosensor Data. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW ’15), Vancouver, BC, Canada, 14–18 March 2015; ACM: New York, NY, USA, 2015; pp. 143–153.

34. Matthews, T.; Rattenbury, T.; Carter, S. Defining, designing, and evaluating peripheral displays: An analysis using activity theory. Hum. Comput. Interact. 2007, 22, 221–261.
64. Intille, S.S. Change blind information display for ubiquitous computing environments. In Proceedings of the Ubicomp ’02, Goteborg, Sweden, 29 September–1 October 2002; pp. 91–106.

65. Miller, T.; Stasko, J. Artistically conveying information with the InfoCanvas: A highly personalized, elegant awareness display. In Proceedings of the Working Conference on Advanced Visual Interfaces (AVI ’02), Trento, Italy, 22–24 May 2002; Association for Computing Machinery: New York, NY, USA, 2002.

66. Stasko, J.; McColgin, D.; Miller, T.; Plaue, C.; Pousman, Z. Evaluating the InfoCanvas Peripheral Awareness System: A Longitudinal In-Situ Study; Technical Report GIT-GVU-05-08; GVU Center Technical Reports. 2005. Available online: https://smartech.gatech.edu/handle/1853/5958 (accessed on 3 December 2020).

67. Chang, A.; Resner, B.; Korner, B.; Wang, X. Lumitouch: An emotional communication device. In Proceedings of the CHI’01 Extended Abstracts on Human Factors in Computing Systems, Seattle, WA, USA, 31 March–5 April 2001; Association for Computing Machinery: New York, NY, USA; pp. 371–372.

68. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. In Proceedings of the 1996 IEEE Symposium on Visual Languages, Boulder, CO, USA, 3–6 September 1996; pp. 336–343.

69. Rhodes, J.S.; Benoit, G.E.; Payne, D.G. Factors affecting memory for dynamically changing system parameters: Implications for interface design. In Proceedings of the IEA 2000/HFES 2000 Congress, San Diego, CA, USA, 30 July–4 August 2000; pp. 284–285.

70. Hartson, H.; Andre, T.; Williges, R. Criteria for Evaluating Usability Evaluation Methods. Int. J. Hum. Comput. Interact. 2001, 13, 373–410. [CrossRef]

71. Pousman, Z.; Stasko, J.T. A Taxonomy of Ambient Information Systems: Four Pattern of Design. In Proceedings of the Working Conference on Advanced Visual Interfaces AVI 2006, Venice, Italy, 23–26 May 2006; Celentano, A., Mussio, P., Eds.; ACM Press: New York, NY, USA, 2006; pp. 67–74.

72. Gantz, J.; Boyd, A.; Dowling, S. Cutting the Clutter: Tackling Information Overload at the Source. 2009. Available online: https://www.officeproductnews.net/sites/default/files/imce/Xerox-white-paper-3-25.pdf (accessed on 20 September 2020).