Exploring Affective Design for Physical Controls

Colin Swindells[*], **Karon E. MacLean**[*], **Kellogg S. Booth**[*], **Michael J. Meitner**[‡]

[*] Department of Computer Science
University of British Columbia
Vancouver, BC Canada V6T 1Z4
{ swindell, maclean, ksbooth }@cs.ubc.ca

[‡] Department of Forest Resource Management
University of British Columbia
Vancouver, BC Canada V6T 1Z4
mike.meitner@ubc.ca

ABSTRACT

Physical controls such as knobs, sliders, and buttons are experiencing a revival as many computing systems progress from personal computing architectures towards ubiquitous computing architectures. We demonstrate a process for measuring and comparing visceral emotional responses of a physical control to performance results of a target acquisition task. In our user study, participants experienced mechanical and rendered friction, inertia, and detent dynamics as they turned a haptic knob towards graphical targets of two different widths and amplitudes. Together, this process and user study provide novel affect- and performance-based design guidance to developers of physical controls for emerging ubiquitous computing environments. Our work bridges extensive human factors work in mechanical systems that peaked in the 1960's, to contemporary trends, with a goal of integrating mechatronic controls into emerging ubiquitous computing systems.

Author Keywords

Haptic display, physical control, design process, affect, rotary Fitts-like task.

ACM Classification Keywords

H5.2 [Information interfaces and presentation (e.g., HCI)]: User Interfaces, Haptic I/O.

OBJECTIVE

Our objective is to understand how the choice of acceleration-, velocity-, and position-dependent force feedback renderings for an active physical control influences user performance. We do this through a controlled experiment that compares the user's performance (measured as response time) and the user's affective (emotional) response (measured both biometrically and by self-report). Two questions are investigated.

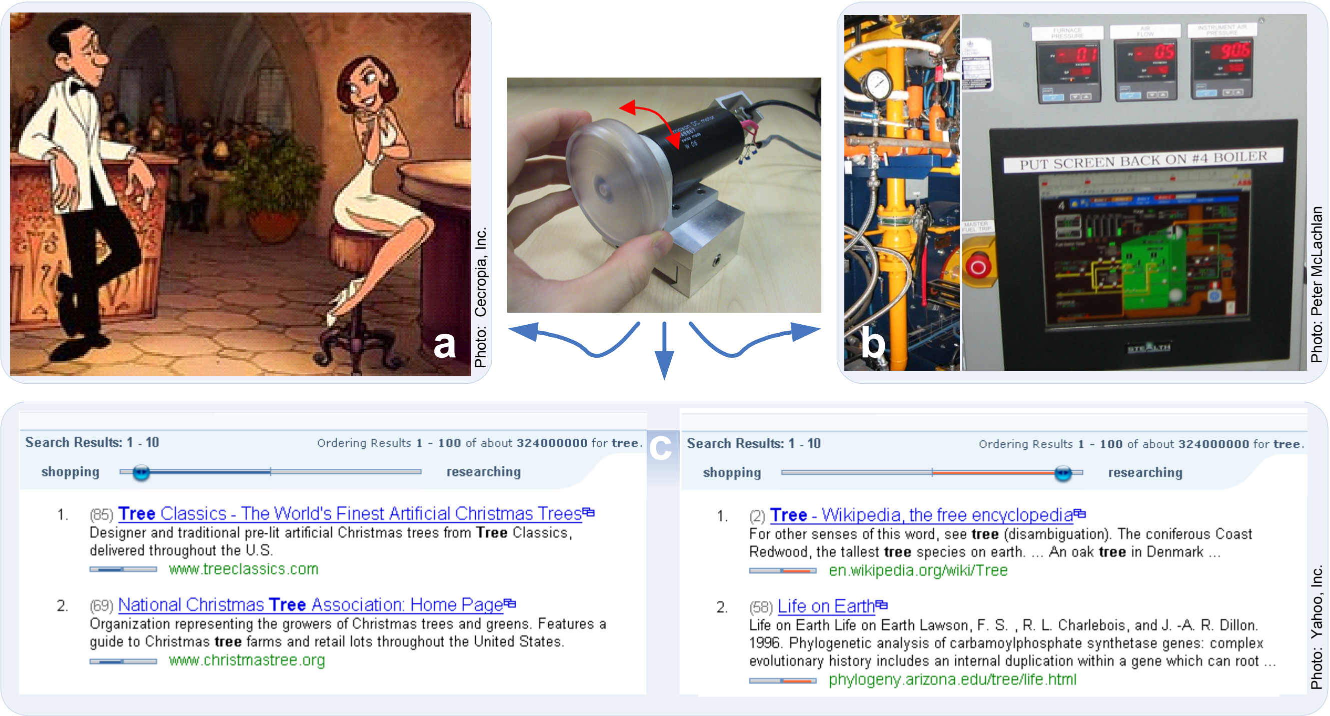
1. How does the 'feel' of moving a physical control influence affective (emotional) responses?
2. How do these affective responses compare with performance when adjusting a physical control?

INTRODUCTION

Mark Weiser notes "the most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" [23]. As we progress from general-purpose computers to ubiquitous, special-purpose embedded computation, keyboards and mice are being replaced by dedicated physical controls such as knobs, sliders, and switches. A new generation of these familiar manual controls are *active*: the way they 'feel' is programmed to reflect measured user actions and situational context. For example, the BMW iDrive, a haptic knob embedded in an automobile cockpit, is designed to help the driver focus more visual attention and cognitive effort on driving instead of interacting with typical 'comfort' features such as climate control and music selection [7]. Because of their pervasiveness in the developed world, it is worth examining our interactions with passive manual control interactions in detail, with the intent of insights for future *active* controls.

In this paper, we focus on *affect* as a potentially potent design dimension for manual controls, because of the intimacy enforced by the need for sustained physical contact and the overall simplicity of these interfaces which highlights what is there. Affective design aspects are already recognized as important in some contexts: as a product line is iteratively refined, its level of adoption by users and its commercial success becomes more dependent on non-technical attributes such as appropriately induced emotional responses [8, 11, 14]. Well-known examples of this are the visceral impact of "heavy" but expensive-feeling stereo volume control knob, and the careful design of the sound and feel of a high-end car door closing or the trademarked throaty roar of a Harley Davidson motorcycle's engine revving. More currently, we observe how customized cell phone cover plates and ring tones (which offer few performance benefits but typically induce strong emotional responses from users) influence sales.

Despite its apparent importance, there is a dearth of mechanisms for actually measuring and utilizing affect in the context of designing interfaces. To address this, we



Search Results: 1 - 10 Ordering Results **1 - 100** of about **324000000** for **tree**.

shopping researching

1. (85) **Tree** Classics - The World's Finest Artificial Christmas Trees
Designer and traditional pre-lit artificial Christmas trees from **Tree** Classics, delivered throughout the U.S.
www.treeclassics.com

2. (69) National Christmas **Tree** Association: Home Page
Organization representing the growers of Christmas trees and greens. Features a guide to Christmas **tree** farms and retail lots throughout the United States.
www.christmastree.org

Search Results: 1 - 10 Ordering Results **1 - 100** of about **324000000** for **tree**.

shopping researching

1. (2) **Tree** - Wikipedia, the free encyclopedia
For other senses of this word, see **tree** (disambiguation). The coniferous Coast Redwood, the tallest **tree** species on earth. ... An oak **tree** in Denmark ...
en.wikipedia.org/wiki/Tree

2. (58) Life on Earth
Life on Earth Life on Earth Lawson, F. S. , R. L. Charlebois, and J. -A. R. Dillon. 1996. Phylogenetic analysis of carbamoylphosphate synthetase genes: complex evolutionary history includes an internal duplication within a gene which can root ...
phylogeny.arizona.edu/tree/life.html

Figure 1: Application scenarios – Different knob dynamics could elicit appropriate user responses for (a) control of character actions in a video game, (b) risk of changing boiler settings in a power plant, and (c) slider bar settings in an on-line search tool.

built a testbed for exploring the relationships between affect (emotional response) and performance. Our first study, reported here, used a rotary manual control (a force-feedback knob), but our approach is applicable to other single degree-of-freedom physical controls such as sliders, switches, and buttons. The experimental procedure measures response time and affective response as users manipulate knobs in a Fitts-like [4] rotational task with varying positions, velocities, and accelerations. The Fitts-task is representative of many 'real-world' applications involving physical controls. Using the testbed, we found that participants generally prefer physical control renderings that improve task performance (we also found counter-examples to this), and we discovered relationships between affective response, task performance, and the parameters of the haptic rendering for the active control that we believe will contribute to successful design guidelines.

In this paper we argue the value of explicit affective design for haptic controls with a series of application scenarios. We then present a user study that compares biometric measures with self reports for measuring affect, and we investigate the relationships of both to task performance. We conclude with a discussion of the new insights for design gained from the study and future steps for research.

EXAMPLES OF APPLICATION CONTEXTS

We first provide three examples to illustrate the broad application space in which affective design of knob dynamics could aid user interaction.

Game Character Control: Figure 1(a) shows a scene from a video game called *The Act* by Cecropia (clips at: http://www.cecropia.com/theAct). Instead of a keyboard and mouse or a complex game pad, character interaction in *The Act* is controlled by a single knob. Rotating the knob could change, for example, the protagonist's level of courage. The central theme of the game is to observe the effects of making characters "charming, tough, sexy, aggressive, sweet, goofy, …" [1]. The knob used in the current prototype of *The Act* is passive (non-actuated). If it were active, its haptic feel could subtly meld with the current emotional context of the video game to enhance a player's gaming experience: the knob could feel 'harsher' as a scene's mood becomes tenser. Determining knob dynamics such as friction, inertia, or detents ('clicks') for a particular emotional context would require an understanding of how these influence a player's emotional response, as well as how transitions between different knob dynamics should integrate with the animated scene, and which dimensions of emotion are most important at a particular segment in the game.

Manual Control of Complex Systems: Many settings for time and safety critical environments, such as the power plant interface illustrated in Figure 1(b), require operators to manually interact with the system either on a routine basis or during emergency override situations. Special cover-plates often act as a barrier preventing accidental use of a sensitive physical control. Adding active feelings to the movements of such physical controls could reinforce safe and unsafe settings to the operator. For example, a knob for controlling atomizing steam pressure or air flow in the power plant could feel 'unpleasant' at risky settings, but feel 'pleasant' at conservative settings. Such psychological reinforcement could be particularly beneficial during emergency situations because reliance on local manual feedback would leave more of the operator's cognitive resources to focus on the emergency situation.

Media Manipulation: Yahoo's *Mindset* is a search engine prototype where users can adjust a slider widget to adjust the content of their query results. With the current user interface (try it: http://mindset.research.yahoo.com), users move their computer mouse to adjust the slider widget towards either "shopping" or "research". Figure 1(c) illustrates results for a search on "tree" – a word with many different context-dependent meanings. Query results from a "shopping" setting focus on Christmas trees or garden stores; whereas, query results from a "research" setting focus on tree biology or computational data structures. If an active haptic control were part of a typical desktop computer setup, an active haptic slider, knob, or scroll wheel could reinforce the current content or subtly provide a wider range of selections. For example, if the search results logically 'chunk' into several logical clusters, the respective number of physical detents could be rendered on the physical control. Further, friction and inertia renderings could subtly suggest previously viewed slider positions or settings believed to be of greater interest to the user by an expert system.

RELATED RESEARCH

There is a rich literature on haptics, including the specific problem of designing knobs, and on techniques for measuring affective response. Our long-term goal is learning how these three areas can provide insights into the design of active control 'handles' for a variety of systems. We briefly summarize key results from the literature that have informed our current work.

Interacting with dedicated controllers

Rogers et al. [18] note that handles such as knobs and mice afford precise or continuous tasks, allow for control-display ratio adjustment, and give (passive) haptic feedback, whereas other handles, such as touch screens, afford direct hand-eye coordination, reduce space requirements, and often require less training and memorization. As we progress towards Weiser's vision of ubiquitous computing seamlessly embedded into our surroundings [23], special-purpose devices such as embedded knobs with active haptic feedback may become increasingly effective and appropriate as interaction components.

When a computing system is dispersed into the user's environment, the space required for input devices is often less of a problem because the devices are built into the environment. A special-purpose interface typically has fewer functions than a general-purpose interface. Consequently, a properly designed handle should be easier to use in ubiquitous computing contexts. Design difficulties will still arise, however, such as the attempts to overload dozens of functions into a single knob in early versions of the BMW iDrive [7] – subsequent versions further reduce driver distractions and improved overall driver acceptance.

A second opportunity is to spatially couple haptics and graphics. Research by Ware & Rose [22] showed benefits of co-locating a physical control with a graphical instantiation of the target. Following this principle, our design integrates a haptic knob into a graphical display.

Design of Knob Physical Properties

Inspired by the need to design dials for rotary phones that 'felt right', Knowles & Sheridan [9] performed early human factors work comparing several friction and inertia parameters for knobs using physical mass and cable pulley mechanisms in terms of both performance (as was common in that era, e.g. [6, 13, 25]) and subjective responses – which was not as common. For example, they found that participants had difficulty detecting < 15-20% changes in friction & inertia, subjects preferred low friction levels, and that subjects prefer at least a small amount of inertia. Such human factors research is again relevant as embedded mechatronic interfaces, including force-feedback physical controls, become feasible and cost effective. This paper takes inspiration from this visionary early work. We have replaced the purely mechanical setup of Knowles & Sheridan with an actively controlled, and thus more versatile, display, and focus on new measures of affect. Current state-of-the-art force-feedback controls can feel *almost* as good as traditional mechanical controls. However, force-feedback technology is rapidly progressing, and force-feedback controls are much more flexible – both from design and usage perspectives. Comparative user studies, such as ours, are an important first step to leverage good quality and extensive early human factors research to guide development of state-of-the-art and soon-to-be-invented force-feedback controls.

There have also been explicit attempts to design controls to display affective parameters. For example, MacLean [12] demonstrated an active door knob with dynamics and temperature which changed depending on the activity behind the door. Thus, a person could use the door knob handle to 'feel' various current and recently past activity states behind the door, including their emotional content.

Recently, there has been attention to rendering active force feedback for one-degree-of-freedom displays. Novak et al. [15] studied the kinematic properties of rapid hand movements in a knob turning task, and fit rotary hand trajectories to a non-linear mass-spring model of movement similar to the underlying model used within our haptic knob. Hasser & Cutkosky [5] modeled a human hand grasping a haptic knob by fitting to a linear, second-order translational model at the fingertip with single constants for rotational acceleration, velocity, and position. We chose haptic rendering models that closely match these human hand models, to ensure that our apparatus will effectively render convincing acceleration-, velocity-, and position-dependent haptic feedback.

Measuring Affective Response

There is a wealth of research, much of it involving human vision, that is based on early work by Russell et al. [19] and

Lang et al. [10], who studied visceral emotion (meaning emotional responses that were not cognitive) and developed models of visceral emotional responses with orthogonal axes of *valence* and *arousal* – often referred to as an affect grid (see examples in Figure 2). Research predominantly using self-reports to measure participant valence and arousal responses to stimuli has found that these two dimensions each account for ~45-50% of the variability in visceral emotion – visceral emotions are effectively modeled as two independent dimensions, with valence typically slightly more influential than arousal in a participant's total visceral emotional response. Others have studied participant responses to more subtle emotional sub-dimensions. For example, Desmet [3], had particpants make self reports after visually inspecting consumer products.

Winton et al. [24] determined that *skin conductance* (SC) measured on muscles in a participant's index and middle fingers (digitus secundus & digitus medius) varied linearly with arousal ratings such as those used by Lang et al. [10]. This SC test is often referred to as a 'lie detector' when used by police because the sensors pick up increased sweat that occurs from elevated arousal levels when a person lies. Parallel work by Schwartz et al. [20] was performed with *electromyography* (EMG) electrodes applied to a participant's facial muscles (corrugator supercilii and zygomaticus major). Corrugator supercilii muscle tension was found to measure valence slightly more effectively than the zygomaticus major measurement. This EMG test is simply a way of examining a person's facial expressions such as frowning. For example, as a person smiles, certain electrical voltage levels fluctuate in facial muscles as they tense and relax. We measured absolute valence and arousal levels using these EMG and SC tests. The more influential visceral emotional dimension, valence, was also measured using the same 9-point rating scale developed by Lang [10] for the Self-Assessment Manikin. One of our contributions is testing the effectiveness of the SAM for haptic research.

USER STUDY

This section describes the experimental design, results, and analysis for a study that measured task performance and preference relationships for knobs that had seven different active or passive haptic controls (refer to the accompanying video for more details of the experimental apparatus and procedure.) These studies involve knob grasps ranging from whole-hand to 1-finger. Our results, which build on our preliminary work [21] are applicable to mechanical controls and other form factors such as sliders.

Participants

Nineteen paid participants (9 female, 10 male) were individually tested in the study that took approximately one hour to complete. All were right-handed. Their ages ranged from 19-35 years ($M = 23.2$, $SD = 3.6$).

Experimental Design and Setup

We used a factorial design based on 7 *knob renderings* × 2 *graphical target amplitudes* × 2 *graphical target widths*.

Participants sat at a desk in a dimmed experiment room and used the right hand to interact with a haptic knob embedded in a graphical display (Figure 3). We used a TimeSys Linux real-time kernel to control the haptic knob, and a Microsoft Windows XP client for the graphical display. Figure 4 illustrates how a typical participant rotated the haptic knob towards a projected graphical disk while feeling force-feedback rendered through the knob. A third computer controlled a touch pad used to collect self-reports, as well as participant EMG and skin conductance (SC) readings. Participants wore noise canceling headphones that played a waterfall sound with a 'near-Gaussian' audio distribution to mask distracting audio cues from the apparatus.

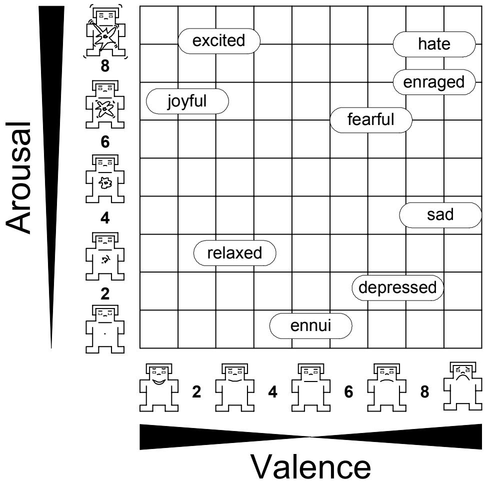


Figure 2: Affect grid of the two most influential visceral emotions: valence and arousal

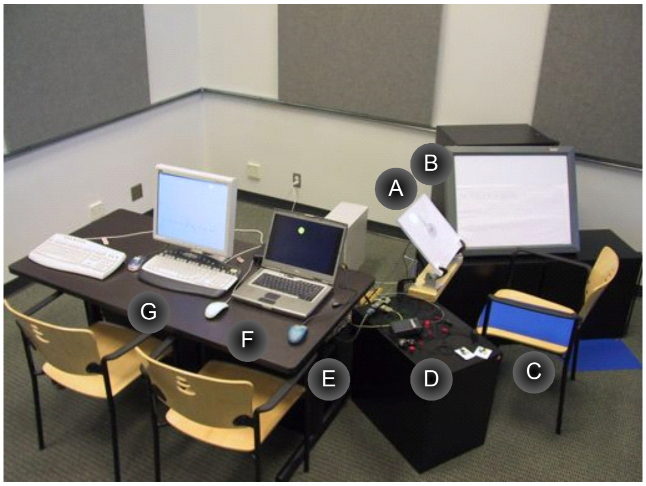


Figure 3: Experimental setup: A) haptic knob embedded into a graphical display, B) touch sensitive surface for self-reports, C) participant chair (with grounding pad), D) biometric sensors, E) trial haptics computer & peripheral hardware (all under table), F) trial graphics scheduling computer, and G) biometric & self-report computer

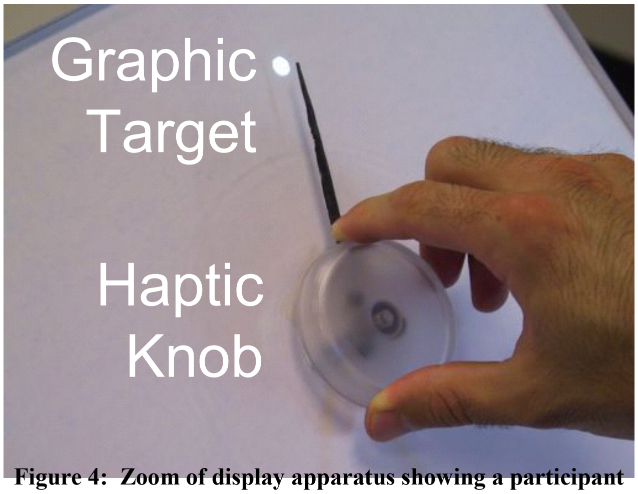


Figure 4: Zoom of display apparatus showing a participant rotating the knob towards a small white graphical target disk while experiencing force-feedback renderings via the knob

Haptic Rendering

Position-, velocity-, and acceleration-dependent renderings were computed according to Equations 1, 2, and 3. Table 1 lists the 7 knob renderings used in our experiment.

$$\tau_{pos} = a_1 \sin(a_2\theta) \quad (1)$$

$$\tau_{vel} = b\dot{\theta} \quad (2)$$

$$\tau_{acc} = m\ddot{\theta} \quad (3)$$

Table 1: Force-feedback knob renderings (Torques in Newton-meters (Nm), times in seconds (s), and angles in radians (rad))

#	*Label*	*b*	*m*	*a_1,a_2*	**Description** ***'Real World' Example***
1	**NON**	0	0	0, 0	No force feedback (control)
2	FR↓	2.6	0	0, 0	Small viscous friction *Portable radio volume knob*
3	FR↑	7.9	0	0, 0	Large viscous friction *High quality sink faucet*
4	MS↓	0	.06	0, 0	Small inertia *Wheel on a small toy car*
5	MS↑	0	.21	0, 0	Large inertia *Fishing reel (free spinning)*
6	DT↓	0	0	7, 180	Compact, subtle detents *Mouse scroll wheel*
7	DT↑	0	0	19, 36	Distant, stronger detents *Box fan settings*

Continuous torques up to 180 mNm were supplied by a Maxon RE40 DC motor. Position was measured with a MicroE optical encoder operating at 640,000 counts/revolution (CPR). A 10 KHz haptic update loop was coded in C++ using RTPM middleware [17]. We custom built this physical setup because haptic knob systems capable of rendering such dynamic effects are not yet commercially available. Typical good quality commercially available systems have encoders operating at 2000 CPR, update rates of 1000 Hz, and continuous maximum torques of 18 mNm (e.g. Immersion's knob for automobiles [7]).

Graphic Rendering

Figure 5 illustrates the knob mounted in the centre of a rear-projected display with 1024×768 resolution and 1500 lumens of brightness. The polycarbonate cap on the knob had a diameter of 64 mm, depth of 13 mm, and a 3 mm filleted edge. The black knob needle was 100 mm in length, extending to the centre of a white graphical target disk that was displayed by the software during trials.

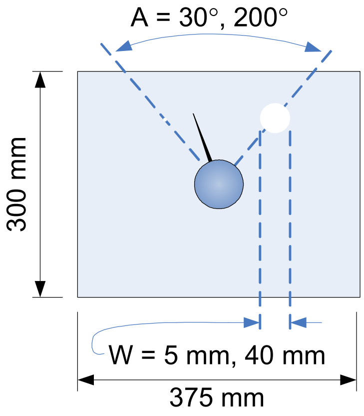


Figure 5: Display apparatus for graphical targets of two amplitudes and two widths. The embedded knob (blue circle in center, with attached needle indicator) is used to point to the graphical target (white circle).

OpenGL was used to drive the graphical display. The graphical client obtained knob position data from the haptics server to maintain a 60 Hz graphical update rate.

Measures

Timestamped data was recorded by the haptic server every 100 µs during each target acquisition.

Self-reports of valence (details below) were measured at the end of each trial using a MERL Diamond Touch touchscreen controlled by a Visual Basic program. A set of nine 3 cm × 3 cm boxes were drawn on the touch screen surface to create a 9-point rating scale where 1 and 9 represented extreme high and low valences, respectively.

EMG was measured at a 32 Hz update rate by placing two AgCl ProComp+ triodes: one centered on the participant's forehead, and one directly above the right eye. The sensors were oriented perpendicular to each other to measure activity of the corrugator supercilii and depressor supercilii muscles, respectively.

Skin conductance (SC) was measured at a 32 Hz update rate by placing AgCl ProComp+ electrodes on the index and middle fingers (digitus secundus & medius) of the left hand.

Procedure

Each participant completed 4 blocks consisting of all 28 combinations of *7 knobs × 2 amplitudes × 2 widths* presented in a different random ordered for each participant in each block. The experimenter read instructions to the participant from a script before the experiment. Participants were given a few minutes to rest between blocks.

Every trial required three rapid movements of the knob back and forth, reminiscent of a classic Fitts tapping task. For each trial, the participant first aligned the knob's pointer over a small, white 5 mm diameter graphical disk. One of the 7 haptic renderings was then applied to the knob. Upon display of one of four possible graphical target disks, the participant moved the knob to acquire it. Once over the graphical target disk, the target disappeared and a second target disk appeared with the *same* traversal distance (Fitts amplitude magnitude) and the *same* diameter (Fitts width) as the first disk, but requiring an *opposite* traversal direction. After rotating the knob towards this second disk, it was replaced by a third graphical target disk of the same size and location as the first disk.

After acquiring the third disk, the participant rated the *appropriateness* of the haptic knob rendering used in the trial for the particular amplitude and target width used in the trial. Specifically, participants were asked to rate how well the current knob rendering helped them perform the last graphical target acquisition. This appropriateness criteria was a more consistent and understandable method for obtaining valence compared to asking participants to explicitly rate 'valence' – a word that many people are not familiar with. Participants were instructed to give a self-report of this valence level on a scale of 1 to 9 by pressing the appropriate graphical cell on the touch pad using the index finger of the right hand.

The three successive target acquisitions in each trial were used to give participants a sufficient amount of time to form a visceral response to each haptic rendering. The repeated angular velocity 'ramp-ups' and 'ramp-downs' as each of the three graphical target were acquired enabled participants to quickly experience consistent velocity and acceleration force-feedback responses. Thus the 'feeling' of each knob rendering was tightly controlled for each graphical target acquisition trial.

The first block of trials was treated as a training task, although participants were not told this. The other blocks were performed to control for three types of apparatus difficulties known *a priori* by the authors. These difficulties were (i) controlling haptic stability during rendering, (ii) maintaining good EMG and SC electrode contact to the participant's skin, and (iii) electrically grounding the response touch pad. Efforts were taken to minimize all of these. For stability, a proportional-derivative-integral haptic torque controller was designed using a root locus technique, the knob velocities were low-pass filtered with a 10th order real-time Butterworth filter, and accelerations were rendered using a 'virtual mass' [2]. To maintain electrode contact, participants were asked to raise their eyebrows and then frown following application of the EMG electrodes.

Biometric responses and, to a lesser degree, self-reports are sensitive to the most minor of experimental disruptions. In an effort to obtain a complete set of high quality data (at the cost of larger data quantities), a block was discarded if the complete apparatus did not perform perfectly for the entire block (e.g., the knob controller had to be stable, the biometric contacts had to be maintained, and the touchpad had to function for every trial in the block). Nine participants experienced at least two blocks with absolutely no disruptions. From these data, the first two blocks containing no disruptions were gathered to form 18 complete sets of data for statistical analysis.

Results

We first tested for data reliability and consistency with previous affect theory. We then examined statistical results to answer our two primary research questions: (i) how do physical control dynamics influence affective responses, and (ii) how do affective responses correlate with physical performance for a given physical control dynamic?

Statistics were performed for the parametric scale measures (SC, EMG, and time) and non-parametric ordinal measure (rating) to achieve two goals: (i) quantify associations between variables, and (ii) compare groups of variables. To quantify associations between parametric and non-parametric measures, the more conservative Spearman correlation was used. To compare groups of three or more parametric groups, repeated ANOVAs were performed, then pairwise comparisons were used to compare individual levels. Similarly, to compare groups of three or more non-parametric groups, a Friedman test was performed, and Wilcoxon tests were used to compare individual levels.

Data Reliability

To validate the reliability and repeatability of our data, we conducted Cronbach alpha standardized item tests on the 18 final cases to ascertain consistency across blocks 2, 3 and 4. This yielded $\alpha = .896$, which is well above the recommended minimum value of $\alpha > .7$ [16]. Data for all three metrics were also checked and confirmed for normality. We concluded that our data were reliable.

Pre-Statistics Filtering

Raw collected biometric data required filtering before statistics could be performed. No filtering was needed for the time and self-report measures.

Single EMG (valence) and SC (arousal) scores for each target acquisition were determined by an independent cognitive science expert using an assessment procedure. To determine a valence score for a trial, the expert observed muscle activity collected from each participant's forehead. Specifically, a raw depressor supercilii EMG voltage trace was subtracted from the corrugator supercillii EMG voltage trace, and the result was low-pass filtered using a 3rd order Butterworth filter with a pass band ripple of 10 dB and stop band attenuation of 40 dB. The expert then manually identified the trial's peak voltage on this smoothed difference trace and subtracted it from the baseline (the flat region preceding a user's knob manipulation) to determine a signed valence value. A positive peak-minus-baseline value indicated positive valence.

The SC technology combines finger muscle voltage measurements into a single low frequency waveform, so no data filtering was needed. To determine an arousal score, the expert manually identified the peak voltage and subtracted it from the baseline to determine an unsigned arousal value where higher values represent higher arousal.

Performance Results

Table 2 summarizes correlations between time and amplitude, width, knob, and rating. Rows show the correlation, ρ, and the level of significance, p. Significant non-parametric correlations were found between target acquisition time and amplitude, knob, and rating.

Table 2: Correlations grouped by time

Time	*Amp.*	*Width*	*Knob*	*Rating*
$\rho(504)$	.123	-.013	-.298	.180
p	.006**	.778	.000**	.000**

* significant at .05 level (2-tailed)
** significant at .01 level (2-tailed)

A repeated measures ANOVA was used to test the amplitude, width, and knob factors. Amplitude and knob main effects were observed for time [$F(1, 8) = 5.3$, $p < .05$, $\eta^2 = .399$, and $F(4.69, 37.5) = 10.9$, $p < .001$, $\eta^2 = .576$, respectively]. A Huynh-Feldt correction of $\varepsilon = .782$ was applied to these knob data to correct for a lack of sphericity. Six pairwise comparisons were performed between (i) the non-rendered control knob and each type of knob rendering, and (ii) the two levels of each knob rendering. Table 3 shows the standard errors (*SE*) and significance level (p) of the time differences for these knob pairs.

Table 3: Pairwise comparisons of time for selected knob renderings

Knob Pairs	FR↑- **NON**	MS↑- **NON**	DT↑- **NON**	FR↑- FR↓	MS↑- MS↓	DT↑- DT↓
SE	.046	.045	.027	.051	.032	.023
p	.040*	.133	.001**	.931	.549	.029*

* significant at .05 level (2-tailed)
** significant at .01 level (2-tailed)

Figure 6 shows the previously described main effects and pairwise comparisons for knob, width, and amplitude vs. target acquisition time.

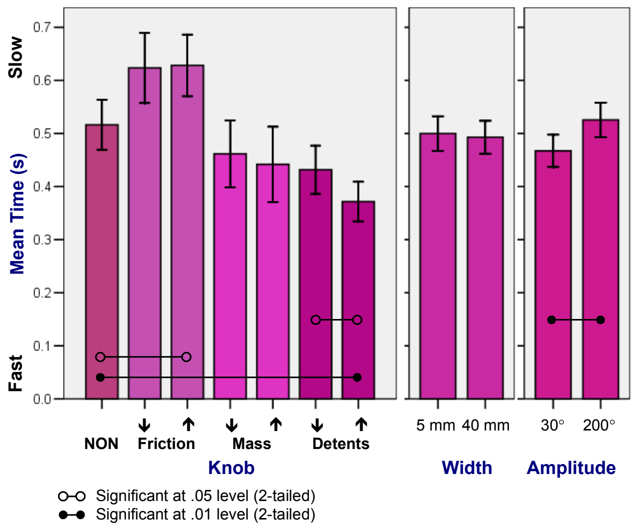


Figure 6: Knob, width, & amplitude vs. acquisition time

Affective Results

Spearman correlations were calculated as shown in Table 4. For each measure *rating*, *SC*, and *EMG*, rows show the non-parametric correlation, ρ, and the level of significance, p. Significant correlations were observed between rating and amplitude, knob, and EMG. Significant correlations were also observed between EMG and knob.

Table 4: Correlations grouped by rating, SC, and EMG

Rating	*Amp.*	*Width*	*Knob*	*SC*	*EMG*
$\rho(504)$	.105	-.034	.119	.053	.135
p	.018*	.443	.008**	.231	.002**
SC	*Amp.*	*Width*	*Knob*		
$\rho(504)$	-.054	-.015	.022		
p	.228	.742	.619		
EMG	*Amp.*	*Width*	*Knob*		
$\rho(504)$	-.005	-.009	.088		
p	.919	.835	.049*		

* significant at .05 level (2-tailed)
** significant at .01 level (2-tailed)

A repeated-measures ANOVA conducted for EMG and SC did not show statistically significant results.

Non-parametric tests for rating showed significant results for amplitude and knob rendering factors. Specifically, Wilcoxon Signed Ranks tests between rating and amplitude were significant [$Z = -3.51$, $p < .001$], and between rating and width were marginally significant [$Z = -1.68$, $p < .092$]. A Friedman test on the knob rendering showed significant rating differences of χ^2 (6, N = 72) = 49.49, $p < .001$. A total of six post hoc Wilcoxon Signed Rank tests were performed on the same six pairwise comparisons that were performed for time (refer to Table 3). Table 5 shows these significant rating differences observed from all six tested pairs of knobs.

Table 5: Pairwise comparisons of rating for selected knob renderings

Knob Pairs	FR↑- **NON**	MS↑- **NON**	DT↑- **NON**	FR↑- FR↓	MS↑- MS↓	DT↑- DT↓
Z	-2.30	-3.37	-3.39	-2.91	-3.32	-4.14
p	$.022^{*}$	$.001^{**}$	$.001^{**}$	$.004^{**}$	$.001^{**}$	$.000^{**}$

* significant at .05 level (2-tailed)
** significant at .01 level (2-tailed)

Figure 7 illustrates the previously described main effects and pairwise comparisons for knob, width, and amplitude vs. rating (valence).

Analysis

Analyses are organized according to four questions.

(1) How did response times (task performance) vary in the knob rendering, target width, and target amplitude levels ?

Figure 6 shows graphic and haptic temporal performance results that one would intuitively expect. Movement times took *longer* a) towards *greater* amplitude targets, b) with *higher* friction knob renderings, c) with *lower* inertia knob renderings, and d) with helpfully spaced detents.

Comparing time with knob NON to renderings FR↓ & FR↑ (Figure 6), higher friction appears to reduce performance. Presumably, the finer control afforded by the additional friction was more than offset (negatively) by the extra physical exertion needed to rotate the knob.

Although only moderately significant (see Table 3), finding similar times for MS↓ & MS↑ that are *both* approximately 10% faster than the control knob NON, despite a 3.5x inertia variation between the two samples, is a helpful finding for haptic designers. Inertia is more difficult to render than friction or detents because accurate acceleration is technically more challenging to measure than velocity or position. Times for these MS↓ & MS↑ knobs suggest that a small amount of inertia improves performance, but larger amounts of inertia provide minimal additional performance benefits for tasks of the sort we studied. Also of interest to designers, times for the detents were similar to inertia renderings, and significantly less than friction and control renderings. Because detent rendering only requires position sensing, programmable as well as mechanical detents are much easier and less expensive to produce compared to inertia renderings. For example, a programmable detent rendering can be designed from a simple potentiometer and braking actuator instead of an optical encoder and servo motor. The shorter times for DT↑ compared to DT↓ are probably due to a combination of (i) high frequency detents more closely resembled continuous friction than low frequency detents, and (ii) 10º / click of DT↑ felt like an intuitive mapping to the 30º and 200º amplitudes whereas the 2º / click of DT↓ felt more like a texture.

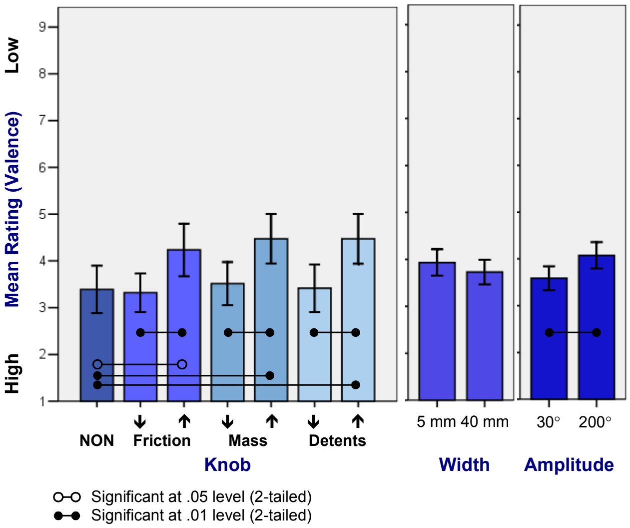


Figure 7: Knob, width, & amplitude vs. rating (valence)

(2) Are the affect results what we would expect?

Finding significant Spearman correlations (see Table 4) between EMG (valence) and rating (valence), but not between SC (arousal) and rating (valence), is exactly what one would expect from previous research [10, 19], which reports the primary emotional dimensions of valence and arousal to be orthogonal.

Main effects for the self-report ratings were observed to be significant even though equivalent main effects for the EMG were not observed to be significant. It is likely that involuntary biometric measurements such as these are calibrated to the full range of human experience, over an individual's lifetime and perhaps over many successive generations of human development (i.e. a full-scale response might be genetically enabled even if never experienced by an individual). Conversely, the self-report ratings for valence span only the context of these stimuli, and subjects are able to voluntarily self-calibrate. Differences between the knob renderings and graphical disks were small compared to levels of previous visual psychology studies such as Lang [10] that compared powerful images including dismembered body parts, furry seals, and nude models. Consequently, the relative significance of the EMG valence to the self-report valence indicate *absolute* valence whereas the individual differences among the self-reports indicate *relative* valence. Thus, our study results suggest that although affectively weak compared to very strong stimuli in other studies, participants could (i) tell the difference between, and (ii) had consistent and measurable preferences for particular position-, velocity-, and acceleration-based knob dynamics.

The remaining analyses focus on these self-report ratings, and utilize the target acquisition times to compare preference and performance relationships.

(3) How did the affective responses vary in the knob rendering, target width, and target amplitude levels?

As shown in Figure 7, many significant self-reported valence differences were observed. Participants generally preferred the more subtle renderings FR↓, MS↓, and DT↓ to the stronger renderings FR↑, MS↑, and DT↑. One might argue that participants were not able to feel the difference between the subtle renderings and the NON knob, but this is unlikely because all the renderings differed in magnitude > 20% from the NON knob as recommended by previous human factors studies using mechanical knobs (e.g., Knowles & Sheridan [9]). The similar valence scores for NON vs. FR↓, MS↓, and DT↓ suggest that haptic position-, velocity-, and acceleration-based renderings can be made to feel as good as a passive mechanical control. This is important because vibrations that occur in virtually all active rendered haptic devices are anecdotally believed to feel unpleasant. The similarity in valence results between NON and renderings FR↓, MS↓, and DT↓ suggest that slight inconsistencies inherent in active haptic controls can be reduced to insignificant levels. The valence findings that generally favor knobs with small amounts of friction (velocity-dependent) and a small amount of inertia (acceleration-dependent) are also consistent with previous findings using mechanical knobs [9]. We thus have further evidence that the friction and inertia renderings used in this experiment correctly model mechanical friction and inertia.

Differences in self-reported valences between the haptic levels FR↓ & FR↑, MS↓ & MS↑, and DT↓ & DT↑, were greater than between the Fitts task-related parameter settings of width and amplitude ($A_{30^{\circ}}$ & $A_{200^{\circ}}$ or W_{5mm} & W_{40mm}). These results suggest that, for this task, haptic rendering had similar or greater effects on the participant valence measures than the pointing task index of difficulties (i.e., different graphical target widths and amplitudes).

(4) How did task performance results compare with affect results?

As a reflection of the complex interdependencies of preference and performance, valence responses sometimes agreed with, and sometimes disagreed with, time responses.

An example disagreement was that participants preferred DT↓ even though DT↑ helped them perform the target acquisitions faster than DT↓ (see Figures 6 & 7). Thus, for tasks where performance is not very important, DT↓ may be a better design selection than DT↑.

Although both amplitude and knob main effects for time were statistically significant, the knob differences may be more practically significant than the amplitude (or width) differences. Mean times for amplitudes of 30º and 200º varied by ~5% whereas mean times for the most extreme knob renderings FR↑ and DT↑ varied by ~25%. These results suggest that designing appropriate haptic feedback for a physical control can influence temporal performance *more* than the spatial organization of the control's settings.

CONCLUSION AND FUTURE WORK

Physical controls such as knobs, sliders, and buttons are an increasingly pervasive and important part of ubiquitous computing systems. This revival of physical controls within contemporary computing systems contains three main differences compared to the pre-personal computer environments prior to the 1980s. First, many information systems have become more sophisticated and complex. Second, contemporary physical controls have improved as a result of better manufacturing processes and greater understanding of human psychophysics. Third, we now have the potential for programmable controls which can respond to a user's context.

As iterative refinements create more mature computing systems, non-technical affective attributes become increasingly important indicators of a system's adoption by users. Like modest performance improvements, modest affective improvements integrate over one's life experience. Thus, relatively small design improvements can add up to significantly improve the overall user experience. Affective and performance responses to a user interface are sometimes correlated, and sometimes not correlated. Consequently, analyzing both affective and performance measures together is crucial for good design. For example, two product enhancements that produce similar performance improvements may induce very different affective responses to their target audiences. A performance improvement that induces extremely negative affective responses will typically result in poor adoption rates. Furthermore, as suggested by Norman [14], situations often occur where people will trade-off product performance if it induces an improved emotional response.

Our contribution is two-fold. First, we have demonstrated the effectiveness of a general process using self-reports and biometrics for measuring relative and absolute levels of the affect induced by physical controls, and we have compared affect and time. Second, we used a validated mechanism to measure affect valence to demonstrate that physical control renderings of position-, velocity-, and acceleration-based effects can significantly influence affective responses. Rendered parameters of the physical knob model were also shown to significantly influence target acquisition times in a tightly controlled performance task; and significant relationships between affective responses and these performance results were discovered. For example, we observed that smaller magnitude knob renderings of friction and inertia were preferred to larger ones, detents that were perceived as textures were preferred to 'louder' more distinct detents, and renderings could be made to feel as good as 'real' mechanical knobs.

Future work should include tests with other haptic physical controls such as sliders and buttons, and different types of tasks. Comparing larger sets of mechanical and rendered mechatronic controls could yield additional interesting insights into the cost-benefit tradeoffs of various position-, velocity-, and acceleration-based dynamics. Instead of adding one dynamic effect to a base physical control, combinations of position-, velocity-, and acceleration-based dynamics could be rendered to better understand relationships between various dynamic properties. Now that we have shown self-reports for valence to accurately reflect biometric data, a similar experiment with a 2-D affect grid could provide further insights into relationships between valence and arousal. Subtle affective attributes represented as sub-regions on the affect grid could then be compared to extensive vision-based studies using the affect grid.

ACKNOWLEDGEMENTS

We thank our sponsors, NSERC, Precarn, and Immersion, and the HCI folks at the University of British Columbia

REFERENCES

1. Adams, E. The designer's notebook: The act – emotion control with single knob gameplay. *Gamasutra*, 12 April 2005.
2. Colgate, J.E., & Schenkel, G. Passivity of a class of sampled-data systems: Application to haptic interfaces. In *Proc. of American Control Conference*, IEEE Press (1994).
3. Desmet, P.M.A. Measuring emotion: development and application of an instrument to measure emotional responses to products. In *Funology: from usability to enjoyment, 3*, Dordrecht: Kluwer (2003), 111-123.
4. Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), APA Press (1954), 381-391.
5. Hasser, C.J. & Cutkosky, M.R. System identification of the human hand grasping a haptic knob. In *Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS),* IEEE Press (2002).
6. Howland, D. & Noble, M.E. The effect of physical constants of a control on tracking performance. Journal of Experimental Psychology, 46, APA Press (1953), 353-360.
7. Immersion, Rotary haptic knob. http://www.immersion.com/automotive/products/
8. Jordan, P. *Designing Pleasurable Products: An Introduction to the New Human Factors*, CRC Press (2000).
9. Knowles, W.B. & Sheridan, T.B. The "Feel" of Rotary Controls: Friction and Inertia. *Human Factors, 8*, Ergoweb (1966), 209-215.
10. Lang, P.J. *The Cognitive Psychophysiology of Emotion: Anxiety and the Anxiety Disorders*. Hillside, NJ: Laurence Erlbaum (1985).
11. Lindgaard, G. & Whitfield, T. Integrating aesthetics within an evolutionary and psychological framework. *Theoretical Issues in Ergonomics Science, 5*(1), Taylor & Francis (2004), 73-90.
12. MacLean, K.E. & Roderick, J.B. Smart Tangible Displays in the Everyday World: a Haptic Door Knob, In *Proc. of HAPTICS,* IEEE/ASME Press (1999).
13. Morgan, C.T., Cook, J.C., Chapanis, A., & Lund, M.W. *Human Engineering Guide to Equipment Design*, (1963)
14. Norman, D.A. *Emotional Design: Why We Love (or Hate) Everyday Things*, New York: Basic Books (2003).
15. Novak, K.E., Miller, L.E., & Houk, J.C. Kinematic properties of rapid hand movements in a knob turning task. *Experimental Brain Research, 132*(4), Springer (2000), 419-433.
16. Nunnally, J.C. *Psychometric theory, 2nd Ed.* New York: McGraw-Hill (1978).
17. Pava, G., & MacLean, K. Real Time Platform Middleware for Transparent Prototyping of Haptic Applications. In *Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS)*, IEEE Press (2004).
18. Rogers, W.A., Fisk, A.D., McLaughlin, A.C., Pak, R. Touch a Screen or Turn a Knob: Choosing the Best Device for the Job. *Human Factors, 47*(2), Ergoweb (2005), 271-288.
19. Russell, J.A., Weiss, A. & Mendelsohn, G.A. Affect Grid: A Single-Item Scale of Pleasure and Arousal. *Journal of Personality and Social Psychology, 57*(3), APA Press (1989), 493-502.
20. Schwartz, G.E., Fair, P.L., Salt, P., Mandel, M.R., & Klerman, G.L. Facial muscle patterning to affective imagery in depressed and nondepressed subjects. *Science, 192*, AAAS Press (1976), 489-491.
21. Swindells, C., MacLean, K.E., Booth, K.S., & Meitner, M. A Case-Study of Affect Measurement Tools for Physical User Interface Design. In *Proc. of Graphics Interface*, Canadian Human-Computer Communications Society (2006).
22. Ware, C. & Rose, J. Rotating Virtual Objects with Real Handles. *Transactions on Computer-Human Interaction, 6*(2), ACM Press (2000), 162-180.
23. Weiser, M. The computer for the 21st century, *Scientific American, 265*(3), 66-75.
24. Winton, W., Putnam, L, & Kraus R. Facial and autonomic manifestations of the dimensional structure of emotion. *Experimental Social Psychology, 20*, Elsevier Press (1984), 195-216.
25. Woodruff, B. & Helson, H. *Torque Sensitivity as a Function of Knob Radius and Load, 80*(4), Univ. of Illinois Press (1967), 558-571.