TABLE 1
The Subject's Descriptions of Imagery and Emotional Character Used for Each of the Eight Emotions

Emotion	Imagery	Description	Arousal	Valence
(N)o emotion	blank paper, typewriter	boredom, vacancy	low	neut.
(A)nger	people who arouse rage	desire to fight	very high	very neg.
(H)ate	injustice, cruelty	passive anger	low	neg.
(G)rief	deformed child, loss of mother	loss, sadness	high	neg.
(P)latonic love	family, summer	happiness, peace	low	pos.
Romantic (L)ove	romantic encounters	excitement, lust	very high	pos.
(J)oy	The music "Ode to Joy"	uplifting happiness	med. high	pos.
(R)everence	church, prayer	calm, peace	very low	neut.

Our work is less concerned with finding a set of "basic" emotions and more concerned with giving computers the ability to recognize whatever affective states might be relevant in a personalized human-computer interaction. The ideal states for a computer to recognize will depend on the application. For example, in a learning-tutor application, detecting expressions of curiosity, boredom, and frustration may be more relevant than detecting emotions on the theorists' "basic" lists.

Clynes' set of eight was motivated by considering emotions that have been communicated through centuries of musical performance on several continents. We started with his set not because we think it is the best for computer-human interaction (such a set is likely to vary with computer applications—entertainment, business, socializing, etc.), but rather because this set together with its method for elicitation had shown an ability to help subjects reliably feel the emotions and had shown repeatable signs of physical differentiation in how subjects' finger pressure applied to a finger rest differs with each emotion [33], [35], a measurable outcome that suggests different states were being achieved. It was important to our investment in long-term data collection that we have a reliable method of helping the user repeatedly generate distinct emotional states.

For the purposes of this research, the specific emotions and their definitions are not as important as the fact that 1) the subject could relate to the named emotion in a consistent, specific, and personal way and 2) the emotion categories span a range of high and low *arousal* and positive and negative *valence*. These two dimensions are believed to be the most important dimensions for categorizing emotions [36] and continue to be used for describing emotions that arise in many contexts, including recent efforts to categorize emotions arising when people look at imagery [37]. The arousal axis ranges from calm and peaceful to active and excited, while the valence axis ranges from negative (displeasing) to positive (pleasing).

2.3 Experimental Method and Construction of Data Sets

Data were gathered from four sensors: a triode electromyogram ($\mathcal{E}$) measuring facial muscle tension along the masseter (with Ag-AgCl electrodes of size 11mm each and 10-20 high-conductivity gel), a photoplethysmyograph measuring blood volume pressure ($\mathcal{B}$) placed on the tip of the ring finger of the left hand, a skin conductance ($\mathcal{S}$) sensor measuring electrodermal activity from the middle of the three segments of the index and middle fingers on the palm-side of the left hand (with 11mm Ag-AgCl electrodes and K-Y Jelly used for low-conductivity gel), and a Hall effect respiration sensor ($\mathcal{R}$) placed around the diaphragm. The left hand was held still throughout data collection and the subject was seated and relatively motionless except for small pressure changes she applied with her right hand to the finger rest. Sensors and sampling were provided by the Thought Technologies ProComp unit, chosen because the unit is small enough to attach to a wearable computer and offers eight optically isolated channels for recording. Signals were sampled at 20 Hz.[2] The ProComp automatically computed the heart rate ($\mathcal{H}$) as a function of the inter-beat intervals of the blood volume pressure, $\mathcal{B}$. More details on this system and on our methodology are available [38].

Each day's session lasted around 25 minutes, resulting in around 28 to 33 thousand samples per physiological signal, with each different emotion segment being around two to five thousand samples long, due to the variation built into the Clynes method of eliciting the emotional states [33]. Eight signal segments of the raw data (2,000 samples each) from Data Set I are shown in Fig. 1. On roughly a third of the 30 days for which we collected data, either one or more sensors failed during some portion of the 25-minute experiment because an electrode came loose or one or more channels failed to sample and save some of the data properly. From the complete or nearly-complete sessions, we constructed two overlapping Data Sets.

Data Set I was assembled before the 30 days were over, and was formed as follows: Data segments of 2,000 samples (100 seconds) in length were taken from each of the signals $\mathcal{E}, \mathcal{B}, \mathcal{G}$, and $\mathcal{R}$ for each of the eight emotions, on each of 19 days where there were no failures in these segments of data collection. The 2,000 samples were taken from the end of each emotion segment to avoid the transitional onset where the subject was prompted to move to the next emotion. A 20th day's data set was created out of a combination of partial records in which some of the sensors had failed.

2. The electromyogram is the only signal for which this sampling rate should have caused aliasing. However, our investigation of the signal showed that it registered a clear response when the jaw was clenched versus relaxed; thus, it was satisfactory for gathering coarse muscle tension information.