H. Knight and R. Simmons. "Laban Head-Motions Convey Robot State: A Call for Robot Body Language." International Conference on Robotics and Automation, 2016.

Laban Head-Motions Convey Robot State: A Call for Robot Body Language

Heather Knight and Reid Simmons, *Members, IEEE*

***Abstract*— Functional robots are an increasing presence in shared human-machine environments. Humans efficiently parse motion expressions, gaining an immediate impression of an agent's current action and state. Past work has shown that motion can effectively reveal a robot's current task objective to bystanders and collaborators. This work investigates whether robots can communicate other aspects of their internal states via motion, e.g., strict adherence to deadline, flexibility of attention, or confidence in a task. Rather than showing us *what* the robot is doing, these motion characteristics leverage the *how* of the task motions to convey additional robot attitudes. To lay the foundations for this objective, we adapt the *Laban Efforts*, a system from dance and acting training in use for over 50 years. We operationalize features representing the Laban Efforts (Time, Space, Weight, and Flow) to the movements of a 2-DOF Nao head and a 4-DOF Keepon robot during simple dance and look-for-someone behaviors. Using online survey, we collect 1028 motion ratings for 72 robot motion videos, achieving significant legibility results for all four Effort implementations. Simple robots may not have human degrees of freedom, but it appears their motion patterns can effectively convey complex expressions.**

I. Introduction

Readable robot motion is essential for human-robot collaboration. A user that can anticipate robot motions will be able to efficiently collaborate [1][2] or move out of the way. *Expressive* robot motion goes a step further, helping convey the inner attitudes of its agent. A robot might appear to complete a task with increasing confidence after many iterations, or with boredom if programmed to seek out novelty. Because of people's acuity at reading non-verbal cues, motion is an efficient channel for encoding expression and can help cue response. Our past work found that varying robot speed impacts whether people will interrupt the robot [3]. Much like a waiter averts his eyes when busy, robots in human-machine environments can use motion to moderate and/or prime interactions with people.

Various robotics researchers have singled out the Laban Efforts as a singularly applicable design framework for robot expressive motion [4][5][6][7]. Many previous adaptations have used a robotic arm or humanoid frame, whereas our work evaluates the Laban Efforts for lower degree of freedom robot head-motions. According to Laban [8], the four Efforts; Time, Space, Weight, and Flow; specify the full space of possible expression for a particular motion goal (e.g., head-yaw moving from 0 to 45 degrees). Thus we use the four polar Laban Efforts as our state space for expressive motion evaluation. To enhance the validity of our results, we evaluate the motions of two robots (see Fig. 1) performing two tasks (see Table I).

*Research supported by National Science Foundation.

H. Knight and R. Simmons are with the Robotics Institute at Carnegie Mellon University, Pittsburgh, PA 15203 (phone: 412-268-3818; fax: 412-268-6436; e-mails: hknight@cs.cmu.edu, reids@cs.cmu.edu).

H. Knight and R. Simmons are with the Robotics Institute at Carnegie Mellon University, Pittsburgh, PA 15203 (phone: 412-268-3818; fax: 412-268-6436; e-mails: hknight@cs.cmu.edu, reids@cs.cmu.edu).

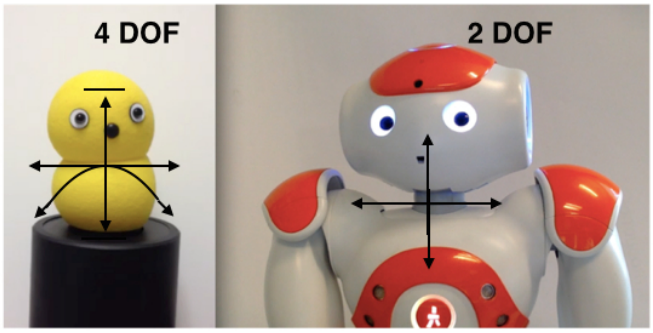


Figure 1. *Illustration of the Degrees-of-Freedom evaluated in our Laban Head-Motions for the Keepon robot (left) and Nao robot (right).*

Each Effort has two poles (see Table II), for example, the Flow Effort can be either *bound* or *free* and the Time Effort can be either *sudden* or *sustained*. We hypothesized that people would be able to distinguish between motion samples with single opposing Effort pole values (e.g., one sample path with *bound* Flow and another with *free* Flow). We provided participants with descriptions of each Effort pole adapted from Laban literature (see Table II), and asked them to correctly label two videos depicting opposing poles of a particular Effort, with all other Efforts held constant. This is non-trivial because there are four Laban Efforts whose features are non-independent (e.g., desired distance might bound the possible speed range) and users were unfamiliar with robot motions [9].

Our online study results substantiate our hypothesis that robot head-motions can convey Laban-defined expressive states, as all four Effort legibility tests show statistically significant results. This is an improvement upon our previous features and pilot [10], in which people had only been able to label the Time and Space Efforts correctly, and we had not yet implemented Flow. Remarkably, we also found that our Effort legibility results were not impacted by robot task or form in the majority of cases. These overall findings suggest that simple variations of robot task motions can allow robots to convey a variety of inner states.

TABLE I. Robot-Task Permutations

	Dance	**Look For Someone**
Nao	nao-dance	nao-look
Keepon	keepon-dance	keepon-look

II. Background & Related Work

The Laban Effort System is part of Laban Movement Analysis (LMA), a system for describing, discussing and documenting human motion, developed in the 1960's to record dance choreography, much like a musical score preserves sound [8]. Laban Motion Analysis has since been used to annotate or explore human movements in dance, drama, nonverbal research, psychology, anthropology, ergonomics, physical therapy, and many other movement-related fields [11]. For example, to gain background in this method, the first author attended a semester-long course in this method targeted at first-year acting majors. The Effort System, which Laban sometimes referred to as "dynamics," attempts to relate interior intention to subtle motion characteristics such as flow and timing.

The Efforts read as if someone had programming in mind (Table II), conveniently specifying a library of combinable motion factors that give insight into an agent's motion. These factors may suggest attributes of the agent's inner state (e.g. confidence, urgency or interest). Laban instructors describe these Efforts as the *how* of a motion. The four Effort factors include utilization of Time: *sudden/sustained*, Space: *direct/indirect*, Weight: *heavy/ light*, and Flow: *bound/free*.

TABLE II. The Laban Efforts

	Fighting Pole	**Inducing Pole**
Time	Sudden: *rushed, hurried, "now, now now!!!"*	Sustained: *lingering, relaxed, waiting for the perfect time to act.*
Space	Direct: *linear, pinpointed, laser-like.*	Indirect: *expansive, flexible, meandering.*
Weight	Heavy: *compressed by gravity, collapsed, overcome.*	Light: *delicate, buoyant, lifted up, floating.*
Flow	Bound: *contained, controlled, rigid, clear.*	Free: *abandoned, released, outpouring, out of control.*

Each Effort runs between two contrasting Effort Poles. Building on Charles Darwin's classification of animal motions into dominant or submissive [10], Laban categorizes *direct, sudden, heavy* and, *bound* as displaying a "fighting disposition," and *indirect, sustained, delicate*, and *free* as displaying an "inducing disposition" [8]. The polarity of each vector indicates the agent's attitude toward that category. For example, an agent's relaxed (*sustained*) attitude toward Time might have gradual velocity transitions, consistent with what one would imagine a deferential animal displaying toward one that is more dominant.

For our application, manifestations of the Laban Efforts need to be adapted to robots. We present a representative sample of computational features (spanning motion classification and generation) that previous researchers have used to represent the four Laban Efforts in Table III. Though we build on previous work in our Effort implementations (see section III), we also make sure to use features that apply to non-anthropomorphic forms. For example, while the inner angle calculations of [4][13] and [14] require a head with torso and arms, overall direction of motion (as in [6][10][15]) is relevant to simple robot forms.

One difficulty in compiling the related work interpreting and generating expressive motion with Laban Effort features is that computational implementations of Laban Efforts are diverse and occasionally contradictory. As one sees in the presented Table III, jerk (the derivative of acceleration) has been alternately used to represent Time, Weight or Flow. Similarly, velocity and/or acceleration features have been used to represent both Time and Weight Efforts. Furthermore, few researchers have implemented Flow.

TABLE III. Laban Effort Implementations by Citation (*indicates Laban expert verification, †our previous work)

	Time	**Space**	**Weight**	**Flow**
[4]†	velocity	orientation	acceleration	range
[5]	rate-commands	inner-angles	kinetic energy	n/a
[6]*	n/a	n/a	acceleration	n/a
[13]	joint-velocity	inner-angles	torque	n/a
[14]	veloc,acc,jerk	inner-angles	n/a	n/a
[15]	slope-of-acceleration	direction-of-velocity	acceleration	n/a
[16]	velocity	area	acceleration	n/a
[17]	duration	horizontal-range	velocity	n/a
[18]	slope-of-velocity	n/a	n/a	n/a
[19]*	slope-of-velocity	n/a	jerk	n/a
[20]*	acceleration	off-axis motion	kinetic energy	jerk

The range of features used to represent each Effort signifies that good classification results alone do not mean that the Laban Effort features were implemented correctly. For example, if a Time feature was used to represent Weight, the system might store the correct information for classification but any interpretation based on Laban motion factors would be wrong.

This disparity is why we provide our study participants with detailed descriptions of the Laban Efforts poles (based on the Table II descriptions [8]), and why we devote this paper to Effort Legibility alone. An alternate approach to verify one's Laban Effort implementation is to use Certified Motion Analysts (Laban experts) to label the motions [6][19][20]. Testing one's implementation across a range of robots and behaviors (as we do in this paper) would further strengthen the generalizability of previous findings using this technique. Laban experts will seldom be the ones interacting with our robots, so it was important for us that untrained users were able to recognize changing Effort values.

III. Our Laban Effort Implementation

While many previous researchers have used ideas from the Laban Effort System or piecemeal features from a particular Laban Effort, our work is one of the first to actively overlay Laban features onto robot tasks. To say this in implementable terms, given a sequence of robot position goals, we use the Effort Settings to set the path and timing characteristics between them such that they might convey a particular (inspired by the animation work of [11]). One of

the most important considerations here is the non-independence of the Laban Effort features (e.g., desired distance might bound the possible speed range).

The non-independence of the Laban Efforts means we need to strategically sequence the application of the Effort features. The two authors worked separately to implement Laban features on the individual robots and arrived at the sample application order (Fig. 4), which is also supported by past work in animation [11]. The constraints on what the goal should be and other feature limits or attributes (e.g., motions should start on beat) come from the task itself.

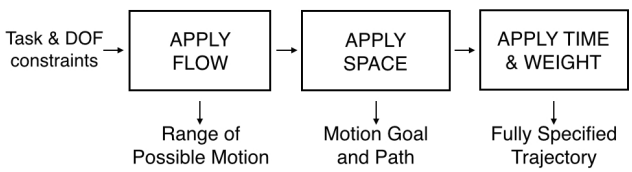


Figure 2. Framework for Laban Effort Application Order

The reason this framework has not been present in the robot literature previous to this paper is simply because people had seldom generated robot motions with all four Laban Efforts, thus they had not yet discovered this ordering. First, we apply motion constraints (mostly Flow), next we calculate path features (Space), and finally we calculate the temporal behaviors along that path (Time and Weight). This approach will become clearer as we review our Laban Effort feature implementations in the sections that follow.

We represent each Laban Effort as a binary variable that has polar values (e.g. the Time Effort can be either *sudden* or *sustained*). One might imagine the Laban Efforts as a four value DIP-switch that researchers can toggle to explore the full state space of Laban expression. Thus, Effort Settings represent the four current Effort Pole values for Space, Time, Weight and Flow (Table II). Extremal values should allow for efficient evaluation of the range of possible motion characteristics and are what actors initially learn to explore the expressive motion possibilities within a scene.

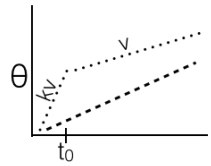
It is important to note that we use the Laban Efforts as the design inspiration for our expressive motion system, but that its features are, by nature, an approximation of the varied way that individual actors might choose to embody these concepts. The essential goal is to evaluate whether simple robot motions appear to communicate inner states during robot task execution, thus our adaptations of the Laban Efforts need only provide a sufficient variety of state expressions for such evaluation.

A. Laban Time Effort

The purpose of the Time Effort is to convey an agent's attitude toward time. In the four robot-task permutations that follow, we have made use of the following feature vector:

$$Time\ Effort = \begin{cases} velocity \\ abruptness \\ arrival\ time \end{cases}$$

Although past works have frequently used velocity to represent the Time Effort (Table III), the idea of a robot reaching its goal before its deadline (*arrival time)* or having a knee-jerk initial response to a new stimulus or general sense



k	2 *sudden*	1 *sustained*
t₀	0.2 *sudden*	n/a

Figure 3. Illustration of Time Effort *abruptness* feature with sample values in seconds, where theta represents rotational motion, and v steady-state velocity. The dotted line presents a *sudden* response with a higher-velocity initial response and the dashed line represents *sustained*.

of jumpiness (*abruptness*) captures additional subtleties of someone who is anxious about temporality. To quantify the latter, we cross-apply time values from the human startle response [21], setting the initial 0.2 seconds of a *sudden* motion traversal to a higher velocity that the rest of the path, as in Fig. 3.

The benefit of maintaining a vector of possible features for each Effort is that one can pick and choose as best applies to the constraints of a particular robot, task or software implementation. For example, the Nao robot does not have the tilt or vertical compression capabilities of the Keepon (when limited to head motions), but it could certainly make use of the yaw and pitch range of motions for Flow, and the acceleration and head pitch features for Weight.

B. Laban Space Effort

The purpose of the Space Effort is to indicate an agent's attitude toward its goal; it could have one clear goal, it could have several possible goals (as in a multi-person conversation) or it might be avoiding its goal or have no particular goal at all.

$$Space\ Effort = \begin{cases} starting\ position \\ target\ Gaussian(s) \end{cases}$$

For simplicity, we have chosen to implement the robot's current goal as a distribution of one (Space=*direct)* or more (Space=*indirect*) Gaussians. The robot can sample this when selecting a new target orientation. We want to control people's first impression of the robots' Space setting, so we hardcode initial yaw orientation values.

C. Laban Weight Effort

The purpose of the Weight Effort is to reflect apparent force. Generally, Weight manifests relative to the agent as either a reflection of the outside forces acting on the agent (*heavy* vs. *light*) or the way in which an agent is actively using force (*strong* vs. *delicate*).

$$Weight\ Effort = \begin{cases} acceleration \\ vertical\ compression \\ head\ pitch \end{cases}$$

We have chosen to use the former representation, as most human versions of the latter involve motions emanating from the core versus from one's fingertips or toes that are difficult to implement on robots with few degrees-of-freedom. Higher accelerations can be an indication of *heavy/strong* Weight. Vertical compression and a downward tilt of the head (Fig. 4, left) are intended to reflect larger outside forces (*heavy*),

whereas vertical extension and uplifted head tilt reflect *light* Weight.

D. Laban Flow Effort

Finally, the purpose of the Flow Effort is to scale all the previous Efforts in a way that conveys the overall constraint or continuity of the agent's inner state. In the literature, interpretations of Flow range from exaggeration (in space) to temporal consistency; we chose to use range of motion.

$$Flow\ Effort = \begin{cases} yaw\ range\ of\ motion \\ pitch\ range\ of\ motion \\ tilt\ range\ of\ motion \end{cases}$$

According to [11], *bound* motions express a quality that the motion could stop quickly if conditions change, whereas free motions display a continuity that would more smoothly transition. We scale the overall extent of motion by varying the robots' rotational ranges of motion (Fig. 4, right).

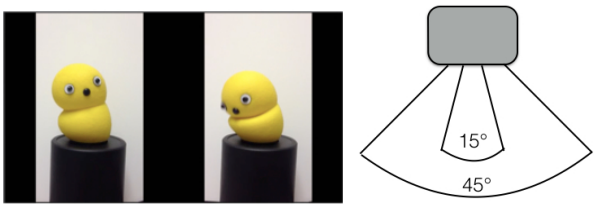


Figure 4. Left image shows Weight Effort *vertical compression* feature on the Keepon robot (*light* left, *heavy* right), Right illustration presents Flow Effort ranges of motion for yaw (narrow range *bound*, wide range *free)*.

E. The Look-For-Someone Behavior

The first task behavior we assess is a *look-for-someone* behavior. Note how we follow the Laban Effort Application Order in Table IV. To accomplish this behavior we sample target positions from a stochastic distribution. This results in a random walk scanning behavior that can move in consistent or inconsistent directions (as depicted in Fig. 5). The motion is centered at the robot's current goal (which we set to be in the center).

TABLE IV. APPLYING EFFORTS TO LOOK-FOR-SOMEONE TASK

Flow	Set range of motion for all joints. *Bound* yaw has a small range ($4\sigma = 15°$), and *Free* has a large range ($4\sigma = 45°$).
Space	Choose starting point and next target position by sampling from range. *Direct* samples from single centered Gaussian (μ=0°), and *Indirect* samples from two offset Gaussians, in which the μ offsets are $\pm 2\sigma$ (calculated from Flow).
Time	1) If Time Effort is *Sudden* and the motion duration allows, have a rapid initial response, at twice the calculated speed, for the first 0.2 seconds (as in human startle response). 2) Apply high velocity cap (1m/s) to *Sudden* to prevent unnaturally jerky motions. Apply lower velocity cap (0.2m/s) to *Sustained* to communicate a relaxed attitude.
Weight	1) Set center point of head tilt to be tilted downward (-0.3 radians) for *Strong,* and uplifted (0.2 radians) for *Light.* Body compression for *Strong*, Body extension for *Light.* 2) Use high acceleration value for *Strong* (0.20m/s^2), lower acceleration value for *Light* (0.05m/ s^2) and the travel distance value from the Space Effort to calculate speed.

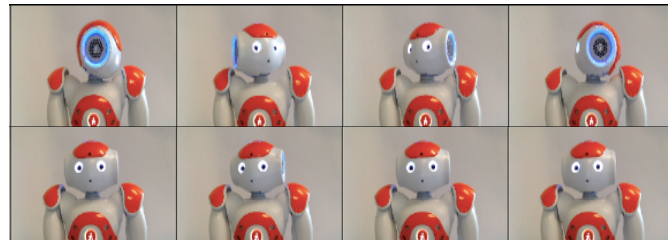


Figure 5. Screenshots from two sample motion sequences. Top row: *indirect* Space, *sudden* Time, *strong* Weight and *free* Flow. Bottom row: *direct* Space, *sudden* Time, *delicate* Weight and *bound* Flow.

Because of software limitations on the Keepon, we were unable to include the first abruptness feature described in Table IV's Time Effort. Because of the lower degrees of freedom in the Nao head, its spatial Weight features include head pitch only, rather than vertical compression.

F. The Dance Behavior

The robots' *dance* behavior had the following constraints: (1) the motions should follow the beat, (2) the motions should oscillate. We review the full procedure for applying each Effort in Table V. We enforce directionality so that the robots move back and forth and bop their heads.

TABLE V. APPLYING EFFORTS TO ROBOT DANCE TASK*

Flow	Set range of motion for all joints. *Bound* yaw has a small range ($4\sigma = 15°$), and *Free* has a large range ($4\sigma = 45°$).
Space	Choose next target position by sampling from Flow distributions. *Direct* samples from single centered Gaussian (μ=0°), and *Indirect* samples from two offset Gaussians, in which the offsets are $\pm 2\sigma$ (calculated from Flow). Enforce directionality constraints, if needed, using absolute values.
Time	Calculate velocity based on distance from Space component and desired tempo (500msec). For *Sudden*, the robot arrives in 70% of the desired period and waits for beat. For *Sustained*, it uses the original metronome value (100%).
Weight	1) Body position for *Strong* is squished (if possible) with the head tilted down. *Light* is less squished and looking up. 2) *Heavy* has a trapezoidal velocity profile to steady state velocity (high acceleration), and *Light* has a triangular velocity profile (low acceleration). We integrate the velocity curve such that we obtain the desired arrival from Time.

IV. LEGIBILITY EVALUATION PROCEDURE

Next, we describe the legibility evaluations of our four Effort implementations across two robots and two tasks on Amazon Mechanical Turk (MTurk) [9]. Our goal is to establish whether untrained users can detect and appropriately label when a particular Effort value is flipped.

To do this, we present MTurk users with two side-by-side videos of the same robot performing the same task in which all Efforts are held constant but one. In the rest of the paper, we refer to these sets of two videos as "comparison videos." For example, in the Time Effort evaluation, the individual videos might both display *direct* Space, *light* Weight and *bound* Flow, but the one on the left has *sudden* Time, while the one on the right has *sustained* Time. We depict two screenshots from a Space Effort comparison video in Fig. 6. It is probably easy to guess which side was more *indirect*.

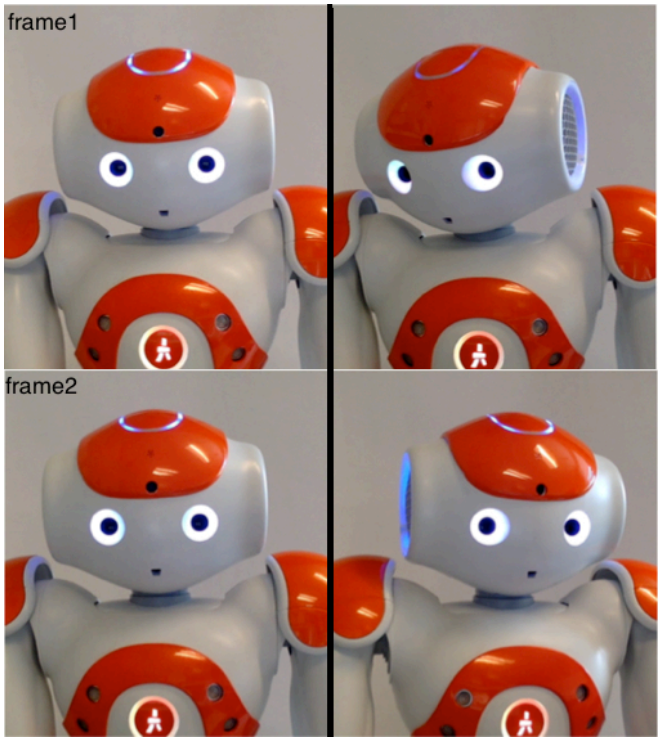


Figure 6. Two sequential screenshots from a Space Effort comparison video of the Nao Look-for-Someone behavior, with *sudden-direct-heavy-bound* on the left and *sudden-indirect-heavy-bound* on the right (**s1**).

We present all possible variations of comparison video orderings in Table VI. To illustrate, in the **t1** videos, the left video has a *sudden* Time setting, while the right video has a *sustained* Time setting. In the **t2** videos, we use the same set of videos with the order reversed, i.e., *sustained* left, *sudden* right. Because there are eight (three factorial) contrasting Effort settings in which one Effort is flipped, and four robot task setups, there are a total of thirty-six videos displaying one ordering (e.g., **t1**), and another thirty-six displaying the reverse ordering (e.g., **t2**).

TABLE VI. COMPARISON VIDEO TYPES (36 CREATED OF EACH TYPE, 8 PER ROBOT-TASK PERMUTATION)

	Left video	**Right video**
t1	sudden	sustained
t2	sustained	sudden
s1	direct	indirect
s2	indirect	direct
w1	heavy	light
w2	light	heavy
f1	bound	free
f2	free	bound

To collect data about whether the MTurk workers can recognize which video has a particular Effort setting, we asked MTurk expert workers to choose which video displayed more of the requested characteristic across all the comparison videos for that Effort (N=72). To do so, we used the survey template and keys outlined below:

> Watch video depicting the motion of two robots side-by-side. Select which robot's motion displayed more of the requested characteristic, or rate the two robots equally.
>
> A- Left more ___
>
> B- Both equally ___
>
> C- Right more ___

**Blank could be:* sudden, direct, bound, heavy, sustained, indirect, free, light.

Before completing the survey, we also provided the workers with a key describing what each rating choice signifies. For example, the full key to the *bound* survey says:

A: *Flow of robot on left noticeably more contained, controlled, rigid, clear.*
B: *Flow of both robots is equally contained, controlled, rigid, clear.*
C: *Flow of robot on right noticeably more contained, controlled, rigid, clear.*

These were the descriptions used for each Effort Pole survey (following format show above):

> TIME1. Sudden: *Timing of robot on left is noticeably more rushed, hurried, "now, now now!!!"*
>
> TIME2. Sustained: *Timing of robot on left is noticeably more lingering, relaxed, waiting for the perfect time to act,*
>
> SPACE1. Direct: *Orientation of robot on right is noticeably more linear, pinpointed, single-focused, or laser-like *relative to CAMERA**
>
> SPACE2. Indirect: *Orientation of robot on right is noticeably more expansive, flexible, meandering, *relative to CAMERA**
>
> WEIGHT1. Heavy: *Weight of robot on left is noticeably more compressed by gravity, collapsed, overcome.*
>
> WEIGHT2. Light: *Weight of robot on left is noticeably more delicate, buoyant, lifted up, floating.*
>
> FLOW1. Bound: *Flow of robot on left is noticeably more contained, controlled, rigid, clear.*
>
> FLOW2. Free: *Flow of robot on left is noticeably more abandoned, released, outpouring, open hearted, out of control.*

For each survey, we collected two ratings for each video from distinct MTurk expert workers to decrease the impact of user error. Specifically, for the *bound* survey, two distinct workers selected ratings for each Flow comparison video (36 of type **f1** and 36 of type **f2**). We also conducted a *free* survey over all Flow comparison videos. We repeated this process for all four Effort categories, running separate surveys to collect data about MTurk ratings for all eight Effort poles.

To review, for each Effort, there are 72 comparison videos (36 in one order and 36 in the reverse), two label surveys and two ratings collected for each. That means there are 72 x 2 x 2 = 256 MTurk ratings collected per Effort and 1024 ratings collected overall. To evaluate our results, we use ANOVA analyses (N=128 because of two MTurk ratings) to test whether comparison video ordering (e.g. *sudden* on left versus *sudden* on right) can predict MTurk ratings. We use a p-value of <=0.05 as our cutoff for statistical significance and a p-value of <=0.1 as a trend toward significance.

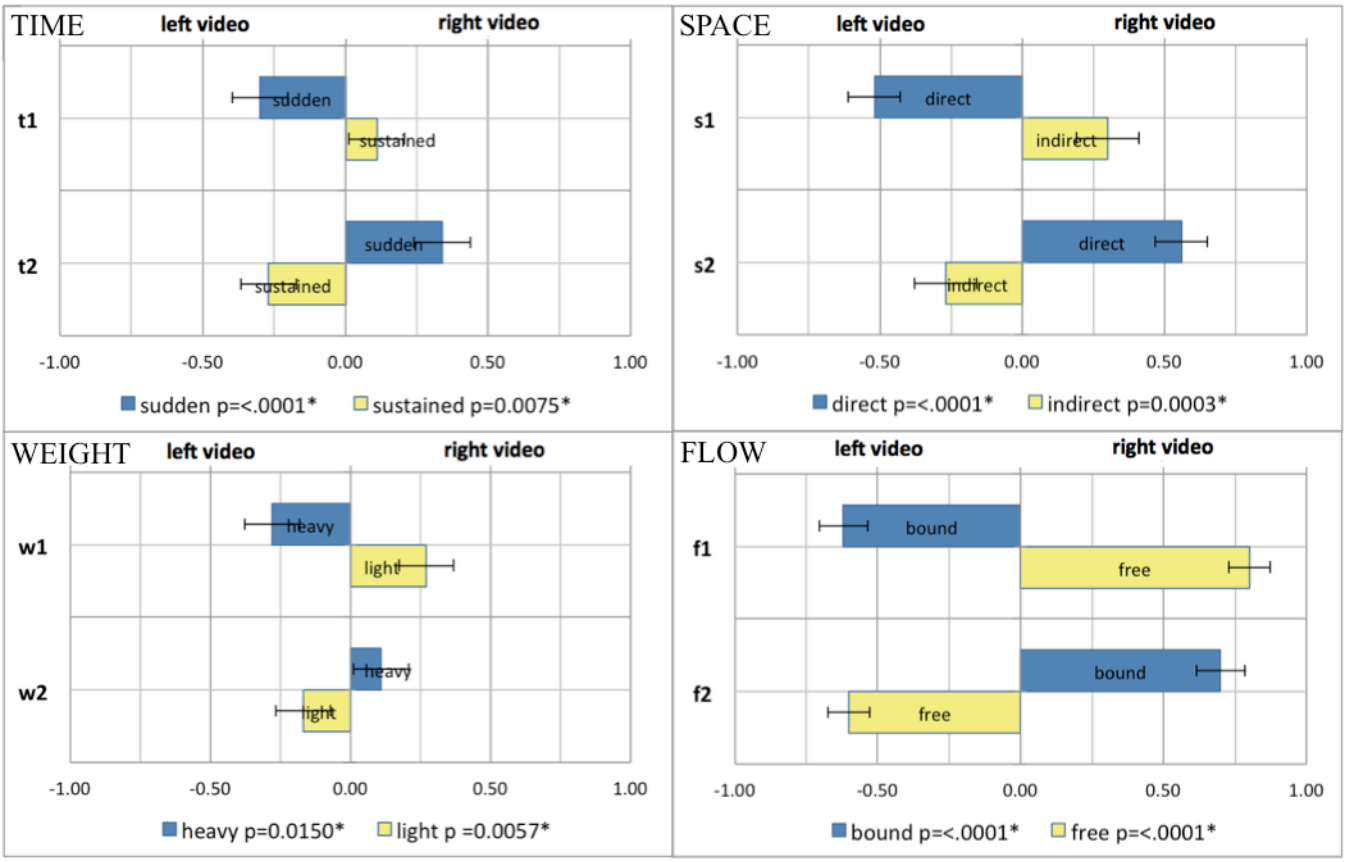


Figure 7. Online Study Legibility Results for Time (*sudden/sustained*), Space (*direct/indirect*), Weight (*heavy/light*) and Flow (*bound/free*) Efforts. The vertical axis presents the two comparison video orderings (e.g. **t1** had *sudden* on the left and *sustained* on the right, **t2** the reverse), the horizontal axis shows mean and standard error of MTurk numerical ratings, and the key presents each label's statistical significance in differentiating between video orderings.

V. Legibility Results

Our approach is to use single ANOVA analyses to evaluate whether video-order (e.g. **f1** vs. **f2)** can predict the mean survey ratings for a particular label (e.g. *bound)*. We find that our Laban Effort implementations have statistically significant legibility ratings for every Effort pole label tested (N=128). Further sub-analyses show that Effort legibility is largely independent of task and form.

We compile the overall results for each Effort by their two labels in Fig. 7, including mean ratings with standard error bars, and p-values linking video ordering to MTurk ratings. Each inset graph in represents one of the four Efforts and includes data from both Effort pole surveys, divided by video ordering. Negative numbers indicate that more workers thought the left video displayed the requested characteristic, while positive numbers indicate more workers thought the right video did. Relating these graphs back to the survey described in section VI, we represent answer A (left video) as -1, answer B (equal) as 0, and answer C (right video) as 1.

We note that mean ratings were always consistent with the true video category (i.e., on average, all videos with *sudden* on the left had a mean rating indicating more people thought *sudden* was on the left). While a mean value of -1 would indicate that every single worker looking at that comparison video type rated the left video as displaying the requested characteristic, naturally, human variation was more diverse with absolute mean values ranging from 0.11 to 0.80 for each comparison video type.

The Flow Effort results were strongest (Fig. 7, bottom-right). We see that the mean ratings of the **f1** videos (*bound* left, *free* right), rate the left video as *bound* (mean = -0.62), and the right video as *free* (mean = 0.80). As one would expect, the ratings reverse in the right set of bar graphs for **f2**. Note that the p-values for both labels are very significant (well below p=0.05, our cutoff for statistical significance).

We see similarly significant overall results for each remaining Effort category. In the top-left Time Effort results, the ordering of the *sudden-sustained* (**t1**) versus *sustained-sudden* (**t2**), reliably predicted the mean ratings of MTurk works for both the *sudden* (p=<.0001) and *sustained* (p=0.0075*) video surveys. Even the lowest absolute means from the overall Weight Effort results could easily distinguish **w1** from **w2** video orderings (p=0.0150 for *heavy* and p=0.0057 for *light*).

Next, we assess the impact of robot form and task. Our hypotheses were 1) that robot form would impact mean ratings and, 2) that robot tasks would obscure the legibility of particular Efforts. Specifically, we expected the dance task to obscure the Time Effort legibility (hitting a beat complicates velocity settings) and the look-for-someone task to obscure the Space Effort legibility (perhaps all scanning behaviors appear to be *indirect*).

	Nao		Keepon	
Time	*sudden* **	*sustained* n/s	*sudden* *	*sustained* **
Space	*direct* **	*indirect* n/s	*direct* **	*indirect* **
Weight	*heavy* **	*light* **	*heavy* n/s	*light* n/s
Flow	*bound* **	*free* **	*bound* **	*free* **

Figure 8. Robot Form legibility results evaluating if comparision video ordering predicts MTurk ratings. ** p<.01, * p<.05, n/s not significant.

	Dance		Look	
Time	*sudden* *	*sustained* n/s	*sudden* **	*sustained* **
Space	*direct* **	*indirect* n/s	*direct* **	*indirect* **
Weight	*heavy* n/s	*light* n/s	*heavy* **	*light* *
Flow	*bound* **	*free* **	*bound* **	*free* **

Figure 9. Robot Task legibility results evaluating if comparision video ordering predicts MTurk ratings. ** p<.01, * p<.05, n/s not significant.

In contrast to these hypotheses, the implemented Efforts were legible to subjects for almost all robot form (Fig. 8) and task (Fig. 9) analyses. We found that 75% of the surveys showed statistically significant results, regardless of either task or form. Note that in dividing our results, we also reduced our analysis sample size by two, so the continued preponderance of significant results is noteworthy.

First, we evaluate the role of robot form (Fig. 8). We display the statistical significance results for each Effort divided by robot form in Fig. 8. We use ANOVA analysis to relate contrasting video order to MTurk rating, this time with N=72, evaluating either the Nao or Keepon ratings for the eight labels. All the robot-Effort pairs have significant results (in at least one direction) except for Keepon weight. There are slight variations in robot means, as per the first hypothesis, one can see that Keepon *sustained* is more strongly significant than Keepon *sudden*, whereas Nao *sudden* legible while *sustained* is not, such the Nao seems more abrupt overall than the Keepon. But most importantly, in the majority of cases, the Efforts are legibility regardless of robot form.

The form results also draw our attention to places where we could improve our Laban feature implementations for each robot. Subjects appear to have difficulty seeing the Nao robot as displaying an *indirect* Space Effort. In this case, we do believe it is the robot form confounding legibility. Perhaps the Keepon's additional degrees of freedoms (e.g. tilt) enact more complex motions that better achieve a sense of *indirect* attitude to its goal. If we were to add more curvature to the Nao motion (i.e. combining the pitch and yaw angles such that the Nao arced into its final goal angle), perhaps we would be able to improve this legibility.

We next evaluate the role of robot task in Fig 9. Again, seven out of eight our task-Effort pairs have at least one significant pole result. We use single ANOVA analysis to relate contrasting video order to ratings (N=72), evaluating either the Dance or Look-for-someone data for all labels. All eight Effort labels showed significance on the Look task, which helps validate our choice of features to represent each Effort for that task category, however we had weaker results for the Dance task, so perhaps the tempo setting did make the robots tend to appear more *sudden* than *sustained*.

The clearer result is that the dance task obscures our Weight Effort features. The act of pushing off the floor or bopping up-and-down that in a common component of dancing motion in general may explain why people do not notice the spatial Weight features (head tilt and vertical compression), and again, the tempo constraints may have also obscured the temporal Weight features (acceleration profile).

VI. Future Work

We challenged ourselves to achieve readable expressive motions with a simple set of degrees of freedom: robot head motions via rotations and compression. Despite this, we have already achieved many significant results.

The next step for the Laban Effort evaluation would be to validate our system in a non-relative context. We wanted to use comparison videos for our first investigation to see if varying simple motion parameters would project robot states to untrained people. Now that we have done that, we would like to evaluate the readability of such robot states in single videos, or more ambitiously, in naturalistic robot settings. The latter would be particularly auspicious for continuing our evaluation in a context where we can evaluate consistent Laban features across a variety of robot tasks.

There also exist many robots that operate in human environments that can traverse space. Thus, the natural next step for us is to extend our Laban motion implementation to mobile robots, adding x-y translation to our possible degrees of freedom. In particular, adding translation could help improve the legibility of the Laban Weight Effort as its temporal representation in our framework is acceleration, and that was difficult to achieve with the short distances traversable by a robot head, and ultimately, we might like to have combinations of Laban Effort such as Space = *direct* and Weight = *strong*, which could, for example, require forward traversal in order to show acceleration if direct was expressed as having a specific heading orientation.

Next, we plan to evaluate the mappings between combinations of Effort settings and specific robot state expressions. We will also explore how they impact or cue human response. In our past work, we found indications that a robot that looks like it is in a rush was interrupted less often

[3]. Given the parameters, there are a large variety of motions and motion combinations we could generate. We are working on interactive tools for learning or exploring such motion parameters in a live setting.

VII. Conclusion

In this work we have operationalized a system from dance and theater training in order to validate the possibility of communicating robot state expressions via robot head motions. We have demonstrated that our system, despite the low degrees of freedom, can reliably convey relative inner states such as hurried vs. relaxed, single-minded vs. flexible, overcome vs. buoyant, and rigid vs. open-hearted.

Although particular researchers have assessed the utility of Laban Efforts to robotics in the past, our work is unique because of the following: (1) we apply the Laban Efforts to low degree of freedom robots, (2) we define features for all four Laban Efforts, and moreover, feature vectors from which motion designers (or an automated algorithm) can pick and choose based on task, (3) we use these features to parameterize robot task motions as opposed to isolated expressions, and (4) we specify an application order for these Effort features that is re-usable across robot forms and tasks.

As we extend this design framework and come up with more ways to evaluate, learn or bring experts in to design expressive motions, we hope to seek out a much larger space of where robot expressive motion could aid a robot's functional or social behaviors. Perhaps people will be more likely to help a robot that physically directs its verbal request at them in a non-aggressive manner. Perhaps we will even be able to use motion features to express a robot's personality (e.g. extroversion/introversion) in ways that effectively customize to a particular user.

Regardless of whether we or other researchers continue to use Laban features, two lessons we have learned during this investigation were: (1) simple robots can convey complex expressive states via motion, and (2) varying robot task motions is sufficient to communicate a variety of expressive states. These results open the doors to using robot body language as a modality of conveying robot inner-states to human bystanders, collaborators and companions.

Acknowledgment

Thanks to the National Science Foundation and Aaron Steinfeld for supporting this work.

References

[1] A. Dragan, K. Lee, and S. Srinivasa. "Legibility and predictability of robot motion." *Human-Robot Interaction, HRI* 2013.

[2] K. Strabala, Kyle, et al. "Learning the communication of intent prior to physical collaboration." *RO-MAN*, 2012.

[3] H.Knight, M. Velsoso and R. Simmons. "Taking Candy from a Robot: Speed Features and Candy Accessibility Predict Human Response." In *Proceedings Robot and Human Communication*, RoMan 2015.

[4] H. Knight and R. Simmons. "Expressive Motion with X, Y and Theta: Laban Effort Features for Mobile Robots." In *Proceedings Robot and Human Communication,* RoMan 2014.

[5] T. Nakata et al. Analysis of Impression of Robot Bodily Expression. *Journal of Robotics and Mechatronics*, 2002.

[6] T. Lourens et al. Communicating emotions and mental states to robots using LMA. *Robotics and Autonomous Systems*, 2010.

[7] M. Sharma, et al. "Communicating affect via flight path: exploring use of the laban effort system for designing affective locomotion paths." In *Proceedings International Conference on Human-Robot Interaction*, 2013.

[8] R. Laban. *Modern Educational Dance.* 1963.

[9] Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. "Using Mechanical Turk as a Subject Recruitment Tool for Experimental Research," 2011.

[10] H. Knight and R. Simmons. "Layering Laban Effort Features on Robot Task Motions," Late Breaking Report at *Int'l Conference on Human-Robot Interaction*, 2015.

[11] D. Chi et al. "The EMOTE Model." SIGGRAPH, 2000.

[12] C. Darwin. *The movements of animals and man.* 1876.

[13] M. Masuda et al. "Emotion Detection from Body Motion of Human Form Robot Based on Laban Movement Analysis." *Principles of Practice in Multi-Agent System*s, 2009.

[14] H. Zacharatos et al. "Emotion Recognition for Exergames using Laban Movement Analysis*." Motion in Gaming*, 2013.

[15] W. H. Kim, et al. "LMA based emotional motion representation using RGB-D camera," *HRI LBR*, 2013.

[16] R. Suzuki et al. "MIDAS : MIC Interactive Dance System." *Conf. on Systems, Man, and Cybernetics*, 2000.

[17] J. Chen et al. "Analysis and Evaluation of Human Movement based on Laban Movement Analysis," *Tamkang Journal*, 2011.

[18] K. Nishimura et al. "Design Support System for Emotional Expression of Robot Partners." *WCCI*, 2012.

[19] J. Kim, J. Seo and D. Kwon. "Application of Effort Parameter to Robot Gesture Motion." *URAI,* 2012

[20] A. Samadani, et al. "Laban Effort and Shape analysis of affective hand and arm movements." *ACII*, 2013.

[21] G. Siegmund , J. Inglis and D. Sanderson , "Startle response of human neck muscles sculpted by readiness to perform ballistic head movements," *Journal of Physiology* (2001), 535.1, 2001, 289–300.