The Four Requirements for Achieving Real-Life Engagement in Online Instructor-Led Training

David Gardner

1 VenueGen, Research Triangle Park, USA

Abstract—Virtual educators know how difficult it is to maintain control over learners’ attention. Without the interpersonal cues and interactions we share when communicating face-to-face, the learning experience can feel unnatural and disengaging, extracting a heavy toll on knowledge transfer. But why do virtual programs often seem less conducive to learning than ILT? The missing element is engagement. This paper will examine:

1. the crucial nature of engagement to successful learning environments;
2. the four factors that create engagement;
3. the ability of various online modalities to engage;
4. how immersive web-3D can help simply and cost-effectively replicate the in-person experience.

Index Terms—e-learning technology, engagement, virtual instructor-led training, immersive e-training

I. INTRODUCTION

Today’s digitally enabled companies are more geographically dispersed than ever before, even as many of their wallets are tightening to fend off economic uncertainty. In an effort to cut costs by moving instructor-led training (ILT) online, the majority of us have at least partially adopted web and video conferencing tools. But in all but the most expensive and elaborate tele-presence facilities (which defeat the purpose of cutting costs), most of these technologies remain inferior to in-person ILT in many important ways.

Instructors know how difficult it is to get and maintain control over learners’ attention with online ILT programs. Without the interpersonal cues and interactions we experience when communicating face-to-face, learners are prone to distraction from a constant stream of more ‘real’ stimuli at their fingertips and on their desktops. They are often preoccupied with figuring out the web conferencing interface, or become frustrated by failed attempts to be heard. If the learning experience feels unnatural or otherwise disengaging—which is so often the case with online ILT—the learner will feel powerless and mentally divided from the instructor and fellow learners. Each of these common symptoms extracts a heavy toll on successful knowledge transfer dramatically diminishing the levels of learner interest, participation, and overall effectiveness of the online ILT programs.

But what exactly is it about face-to-face ILT that seems so much more conducive to learning? Why are retention and compliance scores consistently higher with in-person ILT than with online ILT? And what’s missing from the average online solution that has us consistently wondering whether or not we’re “getting through,” and if the time and money we’re saving by moving online is really worth the loss of effectiveness? The answer to these questions boils down to a single word: Engagement.

In this paper, we will:

1. examine the crucial nature of engagement to successful learning environments;
2. explore the four required factors for creating engagement;
3. rank the ability of various online modalities to foster engagement;
4. examine how fully immersive 3D environments can help trainers simply and cost-effectively replicate the in-person training experience online.

II. ENGAGEMENT: WHAT DOES IT LOOK LIKE?

In any ILT session, whether face-to-face or online, very little can be accomplished until participants feel personally engaged in the learning process.

We know what engagement feels like. When we’re instructing or participating in an engaging activity, we observe that everyone is attentive, thinking, participating, debating and contributing ideas. Learning becomes a collaborative activity. The atmosphere is charged with contagious energy, creativity and a common sense of ownership of the learning process. Simple facts transform into self-discovered truths, and we feel confident that knowledge has not only been grasped, but will be retained and put to use. Mission accomplished!

On the other hand, we also know what disengagement looks and feels like. The instructor might as well be the PowerPoint document from which he or she is lecturing. Discussion is limited to a monologue or is dominated by one or two people, making it difficult for learners to interact, contribute their ideas, or individually engage in the Learning becomes a collaborative activity. The atmosphere is charged with contagious energy, creativity and a common sense of ownership of the learning process. Simple facts transform into self-discovered truths, and we feel confident that knowledge has not only been grasped, but will be retained and put to use. Mission accomplished!

On the other hand, we also know what disengagement looks and feels like. The instructor might as well be the PowerPoint document from which he or she is lecturing. Discussion is limited to a monologue or is dominated by one or two people, making it difficult for learners to interact, contribute their ideas, or individually engage in the Learning becomes a collaborative activity. The atmosphere is charged with contagious energy, creativity and a common sense of ownership of the learning process. Simple facts transform into self-discovered truths, and we feel confident that knowledge has not only been grasped, but will be retained and put to use. Mission accomplished!

The fundamental rule of engagement in group learning dynamics is that people only tend to support ideas, goals, plans and initiatives that they feel they helped to create. In other words, if they feel left out of the process, or as if their input has no bearing on the outcome, then they generally take no ownership of it. Engagement is the power
III. HOW TO CREATE ENGAGEMENT: UNDERSTANDING THE FOUR REQUIREMENTS

So we know what engagement looks like and why it’s important, but how do we create it? It can be difficult enough to build and maintain engagement in an in-person classroom setting, much less online.

To facilitate true engagement in any learning environment—face-to-face or virtual—four components are critically important:

A. Requirement One: The Communication of Focus

Focus communication (Fig. 1) is the ability to convey to others where one’s attention is directed. This characteristic is crucial to the efficient flow of knowledge, and it’s the main reason that in-person training is typically so much more engaging than training conducted online.

In a live classroom setting, we visually communicate that we’re focusing on someone who is speaking by looking in his or her direction. This is significant. It communicates respect and an eagerness to understand the speaker’s contribution. In turn, being at the center of these visual cues triggers a rise in the speaker’s energy and interest levels. Instructors and teachers have known about this phenomenon—termed the spotlight effect—for years. When they move to stand beside a distracted student while lecturing, or ask the opinion of a disengaged learner, they’re doing so because they know how difficult it is to remain a passive bystander while in the spotlight. When all eyes are focused on the disengaged student, his or her energy level will rise, causing a dramatic increase in alertness and participation.

Another important benefit of focus communication involves the complex and fascinating speaking dynamic that occurs when learners interject comments. When several learners begin speaking at once in an in-person ILT environment, some will give way while others get louder as they compete for the floor. The speaker that ultimately gets the floor is determined by a subtle group voting process that happens many times throughout most collaborative training sessions. Learners communicate whom they want to hear speak by offering that person their focus. If the room directs its attention to a competing speaker while we are talking, we recognize the need to relinquish the floor until an appropriate opportunity arises for us to make our points. Focus communication enables group discussion to flow naturally and effectively fostering an engaging interactivity where everyone gets a chance to be heard.

When we cannot visually and audibly communicate our focus to our peers, as when using phone or web conferencing platforms, we tend to talk over one another. As conversation becomes more difficult and less dynamic learners are far more likely to lapse into passivity and distraction. When we do speak while blind to the focus-voting of our peers, we tend to awkwardly talk over each other. The loudest and most assertive person dominates the discussion, often unintentionally frustrating other potential contributors. These would-be speakers often decide it’s not worth the effort to jump in, and the group suffers from fewer, more homogenized collaborative contributions as a result.

From a purely tactical perspective, lack of focus communication makes it difficult to determine who is speaking and to whom comments are directed, putting us at risk of misinterpreting meaning and intent. This is why online learners frequently find themselves wondering whether an instruction was meant for a particular individual or team, or for the entire group. Focus communication also provides important interactive cues to instructors by enabling them to see when learners are looking at the content being discussed, versus when their focus lies elsewhere. This provides important clues to online instructors such as when learners look up from an assignment communicating that they are finished and ready to move on.

With the exception of highly expensive tele-presence systems, all 2D online learning modalities fail to communicate focus. This includes the screen and voice sharing platforms that dominate the current market for web conferencing and online ILT. Even video conference platforms that enable us to see who is speaking provide no directional information such as to whom specifically a comment is being directed.

When focus communication fails to translate through an online learning platform, engagement levels are dramatically reduced. It’s like removing the lubricant from a piece of complex machinery; the gears may still move, but they will be far less efficient, generate more friction and endure significantly greater fatigue.

Here is a simple experiment to illustrate this point: Try blindfolding the participants of your next in-person ILT session. By removing focus communication from your session, you will discover that the group dynamics and participation levels are much more akin to a conference call.

![Figure 1](an exciting and practical advancement in 3D learning is the ability to effortlessly and naturally communicate focus, enabling engagement levels approaching those provided by in-person ILT. Instructors and learners simply click and drag where they want to look, automatically communicating their focus to others. Even eye contact can now be communicated.)
B. Requirement Two: Empowerment

Empowerment (Fig. 2) is the degree to which both learners and instructors feel that they have control within the learning environment. Human nature dictates that when we feel powerless we tend to disengage. Our energy levels dip dramatically, causing us to become bored, passive bystanders who seek stimulation elsewhere.

The ability to choose where we want to look (focus) is very empowering and when we are speaking, we enjoy the spotlight effect and feel engaged but how can we keep learners engaged when they do not have the floor? Wanting to communicate and not being able to do so is not only frustrating, it also creates a feeling of powerlessness that causes disengagement. But, we cannot all be speaking at the same time. The ability to communicate nonverbally helps keep learners engaged between their opportunities to speak.

Beyond the ability to speak and broadcast focus, empowerment also refers to our control over the nonverbal ways we communicate our thoughts and reactions. A simple facial expression or gesture can communicate volumes. During in-person ILT sessions, students and instructors have the power to convey a positive thought with a smile, a negative thought with a frown, and to give quizzical looks when they do not understand something. While listening, we often nod our heads to communicate agreement with a point being made.

These nonverbal measures of control are crucial to engagement because they prevent us from fading into passivity when someone else has the floor. In today’s multi-tasking environment, students need constant outlets of expression in order to remain engaged and active participants. Learners will spend the majority of training time outside of the spotlight—listening, not speaking—which means they aren’t usually under the engagement-boosting effects of center-stage. Being empowered to communicate and express our reactions through nonverbal gestures keeps us actively engaged in learning, even when extended periods of listening are required.

Figure 2: Web conferencing environments based on voice and screen-sharing alone limit participant empowerment. Immersive 3D environments empower students to communicate in ways that keep them actively involved throughout online sessions.

Dr. Tony O’Driscoll’s landmark paper, “Escape from Flatland,” compares two online sessions: one conducted in WebEx®; the other in an immersive 3D modality. The instructor and content were identical, but the learners exhibit vastly different levels of engagement (Table 1).

<table>
<thead>
<tr>
<th>Requirement Three: A Natural Sense of Presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A sense of presence (Fig. 3) is what helps learners feel that they are sharing a physical space together and that they are fully immersed in the learning environment. With in-person ILT, we enjoy a full sense of presence thanks to multiple auditory and visual factors. We are spatially aware of our positional relationships to the room and all people and objects in it. We can see and be seen by those around us in a way that feels natural, real and automatic. We hear someone speaking to our right, turn our heads toward the sound, and immediately recognize Bob as the source of speech. Bob is looking at Amy while he talks, so we know his comments are directed specifically to her. A learner’s sense of presence is tied proportionately to his or her level of engagement. Engagement is fostered by anything that helps free a learner’s mental resources so that the brain can focus as quickly and fully as possible on the training message. The reason for this is physiological as well as psychological. When the ears detect speech, our natural cognitive process identifies the speaker before processing the meaning of what is being said. The brain is tuned to first calculate the physical direction from which sound is coming. This calculation tells us where to look, so we can visually confirm the source of speech. Only then does the brain focus on the meaning of what is being said.</td>
</tr>
</tbody>
</table>
A full sense of presence, like we experience with in-person ILT, makes it easy for our minds to automatically register and process the direction and source of speech so that we can focus undividedly on content.

Unfortunately, most of this rich data doesn’t translate via traditional web and video conferencing. Without spatial and visual information to help us automatically identify the source of speech, we struggle to leapfrog over the missing pieces in order to process the meaning of the words. In essence, because these platforms feel unnatural, they require more of our brain space to navigate and thus detract from our ability to engage in the learning itself. Resistance to this input-scrambled learning process is referred to as audio fatigue.

This seemingly complex concept can be simply demonstrated by imagining for a moment your ILT session from our prior experiment on focus. Now imagine that, in addition to being blindfolded, everyone in attendance can only be heard through a single speaker box in the middle of the room. When both facial images and directional sounds are unavailable, it becomes apparent how much more difficult it is to immediately process, and thus engage in, content. The resulting audio fatigue explains why we feel greater mental strain and tiredness following a long conference call than we do after an equally long face-to-face meeting. It also explains why online ILT programs typically achieve poorer transfer of knowledge than in-person ILT programs.

Because voice and web conferencing platforms provide a weak sense of presence, learners feel that they are simply “listening in” to most online ILT programs rather than engaging in them (which, in most cases, is entirely accurate). Without a shared sense of reality, it feels more acceptable to tune out, perhaps by muting the phone or opening a different web application, to an extent that we would immediately recognize as unacceptable in face-to-face environments.

State-of-the-art 3D learning platforms use a variety of techniques to create a sense of presence that is far more naturally engaging and immersive than typical 2D platforms. The most sophisticated of these 3D environments are equipped with features that provide a sense of presence similar to in-person experiences. For example:

- High-fidelity graphics make the online experience feel real and natural, i.e. participants feel that they are sharing a physical space and objects together.
- Positional sound creates a multi-channel VoIP experience, so users audibly detect speech naturally and realistically from the direction of its source. For example, a learner hears the person sitting to their right through their right speaker, and can then turn to “look” at them.
- State-of-the-art 3D avatars automatically mimic the natural gestures, postures and various expressions that people semi-consciously adopt during in-person encounters. These nonverbal communicators vary to match voice intensity when speaking just like in real life, and seamlessly blend voluntary and involuntary movement to create very natural and realistic experiences.

D. Requirement Four: Personal Intimacy

Intimacy (Fig. 4) is the ability to communicate the human element, creating a more personal learning experience.

Folk wisdom states that if you have something difficult to communicate, you should say it in-person. We are much more congenial face-to-face than we are at a distance. When we cannot see another person’s face, we have a tendency to dehumanize him or her. We are less likely to remember that the object of our frustration, like ourselves, sometimes is in need of the benefit of the doubt and the occasional do-over. Instead, we become more likely to jump to negative conclusions about intent or motive. This same phenomenon that causes road rage during our morning commute puts online ILT programs at risk of fostering impersonal, learning-adverse environments.

That which is more personal is also more universally engaging. In-person training programs are viewed as more personal and enjoyable than online ILT programs, often incorporating team building experiences in which learners can get to know one another. During these sessions, everyone’s humanity is apparent, and cooperation and similarities take a front seat to differences and individual posturing.

---

**Figure 3:** Many instructors are turning to sophisticated immersive 3D environments to create a more natural sense of presence than traditional web and video conferencing can provide. Online learning is no longer limited to shared content. It can now become an immersive shared experience.
2D web conferencing platforms, however, separate our voices from our persons, cultivating an environment that minimizes the human element. Without this element, online ILT often creates less congenial learning environments, thus hampering participation, engagement, and ultimately, learning. It’s for this reason that WebEx® and similar 2D training platforms would not be called upon to perform double duty as team building tools.

Unlike these 2D web conferencing solutions, video conferencing and 3D platforms (that provide photo-created avatars) foster an intimacy level that more closely mirrors physical presence. 3D immersion also affords a unique perspective; with the capability to see ourselves in the third person, we can take a step back and see ourselves the way others see us. Seeing oneself engaged in conversation tends to make us acutely aware of our de-meanors.

IV. How Effective is Each Online ILT Modality at Generating Engagement?

As training programs continue to move online, researchers have defined and studied a multitude of success metrics to compare learning modalities. Studies consistently show that certain immersive ILT environments can create better learning outcomes than their 2D counterparts in the areas of participation, ideation, and retention. However, none of these studies have addressed the reasons for these improvements. With our understanding now of the four requirements for creating engagement in online ILT it is now possible to explain these improved outcomes and to rank each modality in its ability to foster engagement. (Table 2)

<table>
<thead>
<tr>
<th>Modality</th>
<th>Focus Communication</th>
<th>Engagement</th>
<th>Natural Sense of Presence</th>
<th>Personal Intimacy</th>
<th>Total Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td>In person ILT</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Tele-presence</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>3D Immersion</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>Video Conferencing</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Web Conferencing</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

When factors that create engagement are compared, it becomes clear that tele-presence and 3D immersive environments create levels of engagement far closer to that of In Person ILT than video and web conferencing.

V. Engagement is the Key, but What About Other Important Factors?

If engagement were the only criteria we had to consider, then we would conduct all our training in-person. But our time and budgets are limited. As with most business decisions, we make trade-offs in search of the best value inflection points. It is generally best to select the online ILT modality that provides the highest level of engagement while meeting your other criteria and limitations. Following is a comparison of how well each modality meets the most common selection criteria (Table 3).
TABLE III
COMPARISON OF HOW ONLINE E-LEARNING MODALITIES MEET EFFICIENCY CRITERIA

<table>
<thead>
<tr>
<th>Modality</th>
<th>Cost Effectiveness</th>
<th>Reach</th>
<th>Scability</th>
<th>Immediacy</th>
<th>Total Practical Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>In person ILT</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Tele-presence</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3D Immersion</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Video Conferencing</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Web Conferencing</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>20</td>
</tr>
</tbody>
</table>

1= (Not Very Effective) 5= (Extremely Effective)

Physical gatherings (In Person ILT) create the most engagement, but they are also the least efficient. 3D Immersion platforms and Web Conferencing Tools are by far the most efficient, providing the most scalability, reach, accessibility and immediacy at the lowest cost. However, Web Conferencing tools score very low in engagement, while 3D Immersion Platforms score high.

So, although each modality may offer specific benefits to limited use cases, we can conclude that, overall, 3D Immersion platforms are by far the most practical and efficient way to create the highest level of engagement, creating the most value with fewer limitations than other online ILT modalities.

VI. CONCLUSIONS

1. Engagement is the single most important factor contributing to improved outcomes in online ILT training.

2. Engagement is fostered by modalities that enable focus communication, empowerment, a sense of presence and personal intimacy.

3. Different online ILT modalities create varying degrees of engagement.

4. Physical meetings, Tele-presence and 3D Immersion are best at creating engagement, covering multiple use cases with the fewest barriers to use and lowest cost.

REFERENCES


Author

David Gardner. Gardner is CEO at VenueGen, Research Triangle Park, NC 27560 USA (e-mail: dgard-
ner@thevenuenetwork.com).

Manuscript received 01 March 2011.

Published as submitted by the author.