

Designing Stronger AI Personalities

Tanya X. Short

Kitfox Games
Montreal, Quebec
tanyaxshort@gmail.com

Abstract

Chairs' Note: In this invited industry case study, Tanya X. Short introduces eight mechanisms that designers can utilize to better harness procedural character personalities in games. Tanya is co-founder and captain of Kitfox Games, a Montreal-based independent game studio, and a veteran developer and designer known for her expertise in procedural generation and systems-driven game design. With Tarn Adams, she edited the 2017 volume *Procedural Generation in Game Design*. As Tanya explains in this paper, there is an emerging pattern in game design that utilizes character personality as a central gameplay system. Using a number of examples spanning her experiences as both developer and player, in this paper Tanya distills her collected design knowledge into an actionable recipe for building better character personality systems. —James Ryan

Introduction

From a designer perspective, an AI personality is a collection of priorities, algorithms, and behaviours based on environment and context. In fact, I believe there's plenty to be gained from differentiating the way your AI handles the "back-end" of personality (motivations and knowledge) from the "front-end" of personality (relationships and actions/abilities). However, from a player perspective, an AI personality is "what they do and why they want to do it". This may sound similar, but is crucially different, in that players actively want to believe that game characters have real desires and needs, even when they are savvy enough to know it's all code under the hood. To some extent, the nuts and bolts of what knowledge and virtues and vocabulary the character has doesn't matter. What matters is how consistently a character can appear to act according to their nature—consistently enough to be emotionally significant, but not so consistent as to be "robotic".

Thus, the following eight ways to make your AI personalities feel "stronger" to players are all oriented towards guiding the player perception. Having a generated personality isn't enough. Are those personalities interesting? Does the player understand what they are? Are the connections between the reasoning and behaviour satisfying?

In this paper, I'll cover the benefits and risks of player interpretation, secret systems, subtle behaviour patterns, passivity, comedy, after-the-fact investigation, reactions, and transformation.

Guide the Player's Interpretation

Which came first, the motivation or the behaviour? Even if we think someone's reasoning occurs before their actions, we are likely to say Arron usually tells the truth, therefore he is Honest (character judgment) and also say Arron is Honest, therefore he usually tells the truth (logical statement). Both are causes and effects, in their own way. Arron tells the truth, therefore he is Honest, therefore he tells the truth. This is one reason why passive media favors "show don't tell" approaches to storytelling—interpreting others' actions mirrors our real life experiences. This personality-interpretation gameplay is usually found in authored media, such as *Uncharted* or just about any novel.

However, in AI, everything has to start somewhere and most systems do not engage in ad-hoc fuzzy definitions. Causality is key to understanding the meaning of a simulation. Players seek to define a cause and an effect, even if (as is likely) the effect becomes another cause, which creates another effect, and so on. In games like *Crusader Kings II* (Paradox Development Studio 2012) or *Dwarf Fortress* (Adams and Adams 2006; Adams 2017), generated personalities tend to be explained up-front, closer to a "tell then show" approach, with the player taking information about personality traits and comparing it to character behaviour to form a mental model of the underlying algorithm.

In our upcoming game *The Shrouded Isle* (Kitfox Games 2017), villagers are generated with names, portraits, gender, family affiliation, and one random virtue and one random vice. These traits change their ability to contribute to the village, and allow special events to occur. For example, the virtue Accusatory (seen on Nadya in the upper right of Figure 1) mostly means that the character is skilled at Penitence tasks, but also means that the engine may select them to populate certain random encounters, such as ones that involve wild accusations. Importantly, in the screenshot shown in Figure 1, Nadya's "Accusatory" virtue happens to have been previously discovered and revealed to the player. In this case, the player can use their knowledge of Nadya's personality as a factor in how to react to the situation. Knowing



Figure 1: A personality-driven event in *The Shrouded Isle*.

that Nadya is Accusatory, they may be more skeptical of this report—they use their knowledge of her personality to decide whether or not she is telling the truth. If Nadya’s virtue were only displayed as “???????” in a Shrouded Isle event, the accusation event would then provide a clue—maybe Nadya is Accusatory, or maybe Arron really is a blasphemer, or maybe both. Later, when the truth is discovered, this is a learning/“aha!” moment for the player.

In another Kitfox title, *Moon Hunters* (Kitfox Games 2016), we made the traits a reactive system entirely—the game tried to interpret the players’ behaviours and choices into mythic hero traits (Cunning, Foolish, etc.). The behaviours were experienced by the player and used to calculate the “reasoning” behind them, which was then displayed as the player’s personality traits. It’s possible this might have been more satisfying for players if they had been able to input on the reasoning itself, or further explore the influences of one on the other. In order to employ the other tips below, you must first decide which your system and player experience depends on first—does behaviour determine personality or does personality determine behaviour? And how does it appear to the player?

When it comes to physicality, there are many highly effective cues you can use to tap into human biases, assumptions, and instincts, such as the human penchant to detect faces and eyes, or to be alarmed by certain colors, etc. But connections between reasoning and behaviour (i.e. personality) are difficult to “see” and based more in culture than biology. Personalities prove themselves over time, and are subject to interpretation. It’s tempting to hide some reasoning elements (needs, desires, virtues, etc.) because it’s clearly more natural. Behaviour-first messaging makes your characters more life-like and puts behaviour at center stage, allowing for gameplay closer to authored content. But why are you generating personalities in the first place if you could just generate the behaviours directly? In a game without exposed reason-



Figure 2: A screenshot of the starting quest area in *Age of Conan: Hyborian Adventures*.

ing or personality-related gameplay, maybe “random” really is good enough. In the current game landscape, it’s fair for players to assume that personality elements that are invisible are, in fact, missing. Players don’t have a good reason to expect AI to have psychological depth. You might need to be fairly blunt about your systems if you want your generated behaviour to be noticed (and understood correctly).

There are at least three ways in which normally advisable, narratively intriguing subtlety can backfire when implementing personality generation systems: secret systems, subtle behaviour patterns, and passivity.

Beware Secret Systems

I joined the *Age of Conan: Hyborian Adventures* (Funcom 2008) team as an AI Designer about a month after the game’s launch, having been an ardent supporter pre-launch. I was a fairly engaged player in those first weeks, almost reaching the maximum level. However, I was stunned in my first

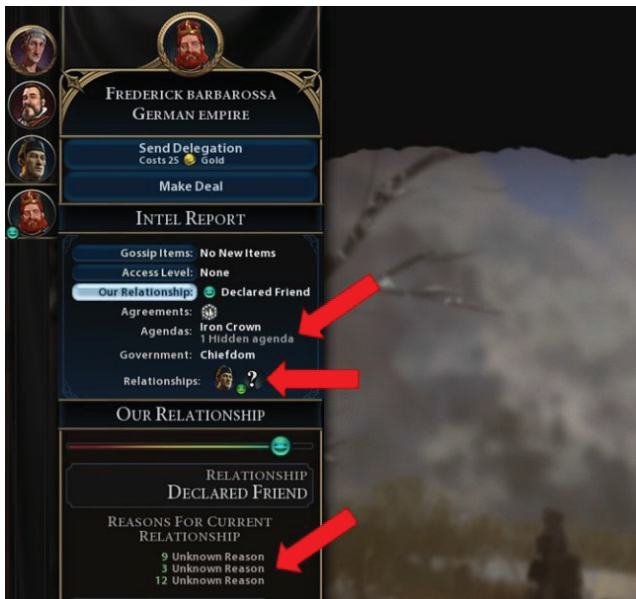


Figure 3: The player-facing details of *Civilization VI* opponent AI status, with arrows pointing to its goals, relationships, and motivations, some of which are tantalizingly obscured.

few days at work to find that almost all NPCs had a deeply modelled motivational system based on the Maslowian hierarchy of needs (Maslow 1943). It was complex, modular, powerful, and completely invisible to players. Even characters whose only purpose was to populate “grinding” zones (masses of creatures spawned only to be killed en masse for currency or experience points) would first seek safety, then to eat food, then to sleep if tired, then to socialize, and finally to self-actualize. It was sophisticated, expressive, modular, and completely invisible.

But if the player doesn’t know trees exist in your game’s forest... can the tree actually fall?

As a player, I had assumed soldiers slept in their tents or talked to one another or walked on patrol because they had been scripted to do so; a “living breathing” world had been constructed for me, and I had missed it! If *Age of Conan* AI had been structured differently or expressed their reasoning (hunger, sleepiness, etc.), maybe I could have detected causality. But the NPCs’ needs grew invisibly, and the core gameplay (combat, combat, combat) gave me no incentive to watch for those expressions, even if they had them. So, learning from *Age of Conan*, I understand why *Dwarf Fortress* exposes as much as it does. It’s the safest route to making sure your effort actually influences the player perception. It might be overkill in some instances, but causality is easier to determine.

At the least, for those elements we hide from the player, it seems safest to hint that they are there somehow. *Civilization* (Firaxis Games 2016) ruler AIs may have become increasingly complex over the years, but they’ve also started being more coy about hidden elements, hinting at them overtly (as shown in Figure 3). This piques curiosity, and prepares



Figure 4: A ruler in *Crusader Kings II*, including motivations and personality traits.

the player for emotional satisfaction when those elements are eventually revealed. Non-subtle obscuring allows the player to engage in both sides of the causality chain, with a bit of personality-interpretation gameplay up front and then personality-prediction gameplay after. I’m interested in the future of *Hello Neighbor* (Dynamic Pixels 2017), which claims you match wits against an advanced, “learning” AI as its central marketing tenet. I’d love to take bets on how subtle or blunt their system communications become over time as they test and refine their player experience... and in this example, the neighbor doesn’t have a procedural personality as such—booby-trapping your basement is much more concrete than expressing your core identity.

Beware Subtle Behaviour Patterns

Extreme personalities are supposed to be rare. If you choose to become inspired by findings from real-life psychology, such as the Big Five personality model (Digman 1990), you’ll see science and realism find the most common traits are something like “mildly confident” or “mostly cooperative”—you know, relatively normal. This is a trap for the same reason that hidden systems are: the player might never actually see someone act in an interesting manner.

However, if you ignore all the ways in which a character is assumed to be ‘average’ and instead focus their behaviour on the limited ways in which each character is extreme, and push it even a little further than realistic, it’s more likely that their traits will be observable for the player. There’s a reason characters in *Crusader Kings II* are Wrath, not “irritable”, and only part of it is to do with poetic archetypes. The other part is that the systems are easier to understand and detect.

Archetypes also tend towards extremes—we’re used to it even in traditional authored storytelling. Ravenclaws aren’t just “kinda smart”, they are defined by being the smartest.

Beware Passivity

Pro-active behaviour is generally less risky to base a personality on than passive or avoidant behaviours. A character that wants to do something is more easily perceived and un-

derstood than a character that wants to not do something, or even worse simply has no relevant desires. This can be a problem for some character types that are natural in storytelling, who are mostly passive, shy, avoidant, or easy-going. Many classic virtues (such as chastity or temperance) were primarily defined by their preference to avoid popular activities. Arguably, the most successful personality type in *King of Dragon Pass* (A Sharp, LLC 1999) was the Trickster, who was the most flamboyant and notable of the characters, while the rest were mostly measured and balanced in their approaches. I asked its designer David Dunham and he says he felt it worked well and will use similar archetypes in his next game, *Six Ages* (A Sharp, LLC 2017). (He also made a few other comments, which I'll get back to in the summary at the end of this paper.)

There are ways to highlight more passive characters, but generally, the more you can re-orient your design to provide opportunities for the character to express their trait actively, the better. For example, rather than merely abstaining from sexual activity, which might take many years to observe, a Chaste character could recoil in horror from sexual content, shame others who engage in flirting, or actively put effort to try and Avoid Temptation. All of this is closer to a detectable pattern for the player, even if a bit silly.

Lean Into the Comedy

In authored media, a good drama can be easier to pull off than a good comedy. Drama can be relatively formulaic; comedy requires continual surprise. Maybe this could be its own case study, but I'm going to assert here that it's the reverse for system-driven "stories": good comedy is easier to pull off in procedural personality systems than good drama. Why? I'm going to assume here that you aren't trying to fool the player into thinking your characters are authored content and I will assume your players know that your characters are generated. As far as I'm concerned, trying to produce Turing-test-passing procedural storytelling is a bit like trying to sell roller skates to someone who wants a bicycle—it completely throws away advantages you might have exploited. Machine collaborators can produce many flavors of narrative and art; if you want an authored story, get an author. So, whether it was part of your marketing or core to your gameplay, *your players are engaging directly with the fact that these characters are system-driven*. Nobody is under the illusion that your characters are people, or even human-authored people! Once players are mentally in the intellectual space of engaging with AI pro-actively, they are emotionally far from the submissive position of suspended disbelief or "immersion".

In order to reach the kind of emotional investment in these little characters that we get from authored stories, players have to internalize your systems completely. Only then can they map humanity onto the characters. That means there's (potentially dozens of) hours in-between when the game starts and when drama can begin. I suspect that *Dwarf Fortress* may never have become popular if its core gameplay didn't start out less personality-driven and more accessible survivalist. By the time someone has created an interesting user story involving their dwarves' personalities and

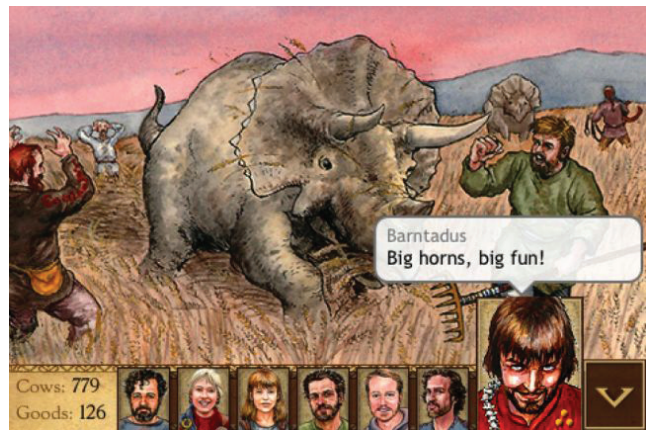


Figure 5: Noticeably bad advice from a Trickster advisor in *King of Dragon Pass*.



Figure 6: Two characters respond differently to stimulus in *The Sims 4*.

eugenics, they are 40+ hours in. In the hours of gameplay leading up to the players' moment of "grokking", it's much easier to make funny situations than it is to make riveting drama. Your characters are the symbols of people, which gives them an awful lot in common with the setup for a joke. Jokes typically use symbolic theoretical people and situations anyway. *A/an (insertnoun) and a/an (insertnoun) walk into a bar...*

Encourage After-the-Fact Investigation

In an ideal world, the player always understands everything going on, and is in fact filled with anticipation for an important, complex event. However, when exploring interactive simulations, especially of the human psyche, it's totally reasonable for a player to ask "Wait, what just happened?"

Assuming there are multiple AIs in your systems (presumably with different personalities and needs and behaviours), players may not be focusing on the right place at the right time. It can then be helpful to provide some kind of tool for the players to play detective and opt into a deeper level of systems can help diffuse what happened and why.

A few example tools:

- Logs or journals of character actions/behaviours.

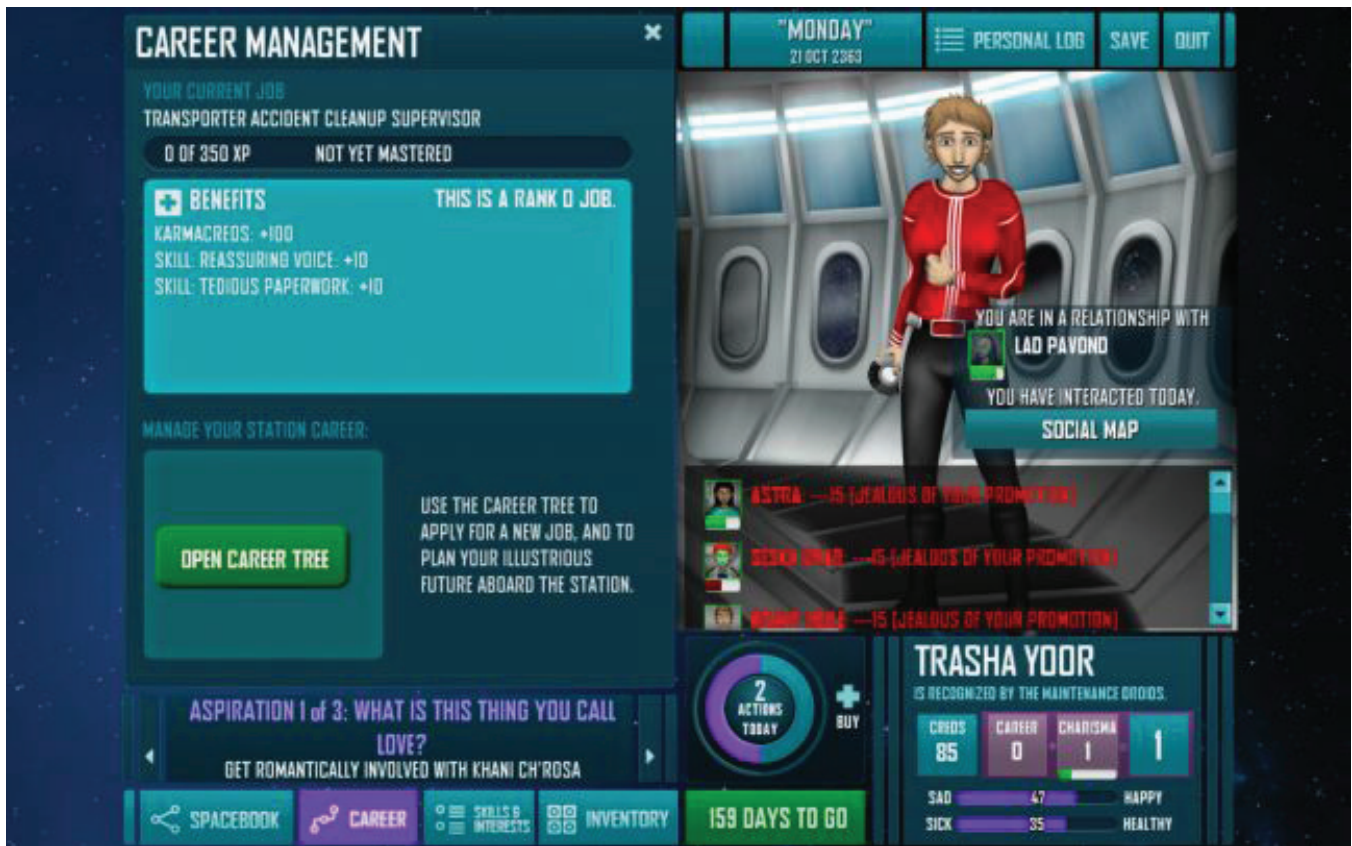


Figure 7: Several crewmates respond with jealousy to your promotion in *Redshirt*.

- Rewind time to re-play and watch events with different actors.
- Actual in-character investigations (ask characters questions, etc.).

This can easily add a whole new dimension of scope to your gameplay features, but as long as you're creating gameplay about personalities, maybe it's what your game needs anyway.

Reactions are Worth More than Actions

As Epictetus allegedly wrote on matters of personality, "It's not what happens to you, but how you react to it that matters" (Epictetus and Higginson 1944). Normally, in order to observe multiple AI personalities, you observe them in sequence following their proactive "natural" inclinations; A tells the truth a lot and B lies a lot, so maybe A is more Honest than B. However, you can process more information per second if you can compare simultaneous character reactions, especially if they're extreme, say to someone's injury or misfortune, such as when one person's tragedy is another's hilarity, as in the scene from *The Sims 4* (Maxis 2014) shown in Figure 6. Or when you can quickly gauge a whole crowd's response to a sweeping change, as in the example from *Redshirt* (The Tiniest Shark 2013) shown in Figure 7.

It's worth a quick warning that part of the reason reac-

tions are so effective and economical is because we have so many expectations surrounding them. Reaction-based personality expressions can become complicated much more quickly than proactive solo actions. The more factors in your system, the harder it is for a character to pick a consistent or even somewhat appropriate reaction. Human priorities are not particularly elegant or obvious. For example, when a character dies, does your AI care more that A) it hated the person who died, or B) that it is a kind person, or C) their judgmental friend is in the room, or D) the person who died did so in an embarrassing way, or E) the murderer was their lover, etc.

Use Transformation (Sparingly)

People (or at least their needs and desires) tend to change over the course of their lives, due to experiences, learnings, traumas, and/or nature. Although it might seem counter-intuitive, having someone's personality change at key moments can actually be more compelling than the personality itself. In the face of adversity (or the ravages of time), some people get weaker/more flawed, a few get stronger/less flawed, and some get weirder. That's how life goes.¹ *Darkest*

¹Though, in Western culture, at least one study shows that people generally tend to become slightly more confident, conscientious, and emotionally stable throughout adulthood, but decrease in

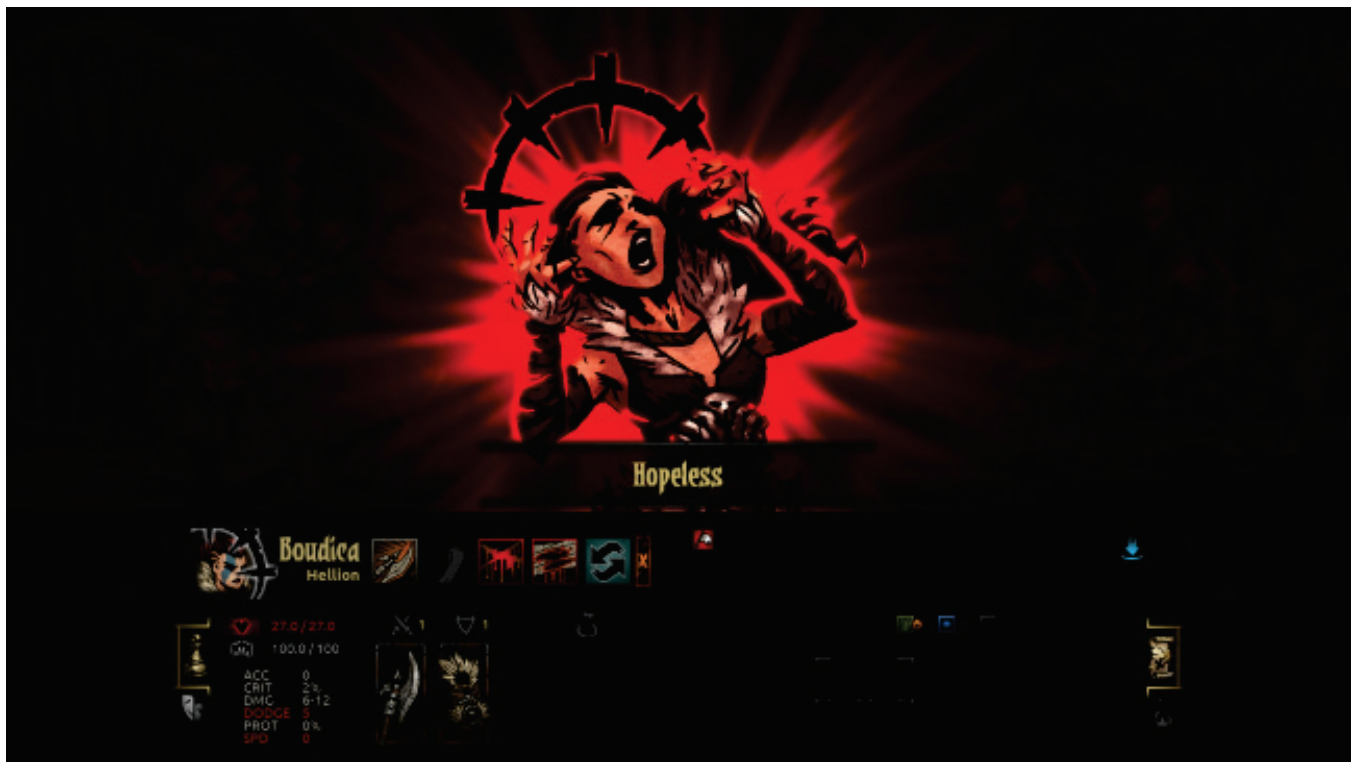


Figure 8: Stress has caused this hellion to become Hopeless in *Darkest Dungeon*.

Dungeon hinges some of its most compelling gameplay moments (the gain and loss of “Quirks”) on moments of crisis and character personality changes. Its designer, Tyler Sigman, confirmed to me that Quirk gains are relatively arbitrary, with some broad exclusions based on mission content.

Crucially, transformation is fascinating on its own as a learning moment with the character and doesn’t need deep explaining of factors and motivations, assuming players are given enough information (state 1, trigger, state 2) to invent their own reasoning.

Summary

Although the RTS and Strategy genres have been dabbling in personality generation for decades, personality as gameplay system is a relatively new subject, without the decades of exploration and discussion of terrain or text generation.

Here’s a quick summary of the insights discussed above:

- **Chicken and Egg.** Have a clear vision for whether reasoning powers motivation or vice-versa, from the engine and from the player perspective.
- **Beware Hiding Too Much.** Be as clear as possible about what’s under the hood.
- **Beware “Normal” Personalities.** Extremes are easier to see and understand.

sociability and openness to experience in older age (Roberts, Walton, and Viechtbauer 2006). But as Shakespeare shows us, compelling narratives need not follow this scientific guideline.

- **Beware Passive/Avoidant Traits.** Actions are easier to see than non-actions.
- **Embrace the Comedy.** Your AI were never human anyway. Sorry.
- **Get Out the Magnifying Glass.** Empowering a little player detective-work can help defray chaos.
- **Reactions Are High-Value.** Two characters + one catalyst = two personalities.
- **Transformations Are Gold.** Even if your traits or behaviours are subtle, a strong and clear change can be compelling on its own.

I’ll also note a counterpoint to this advice: *King of Dragon Pass* is an excellent example of a beautiful, compelling game with procedural personalities that doesn’t use many of my tips—it’s subtle and dramatic, and very successful in achieving its goals. In fact, when I asked David Dunham, he explicitly warned that subtlety can be helpful when trying to tell a convincing story; if you aren’t subtle enough, you could end up with cartoon characters rather than people.

Choose your design risks carefully; if you are willing to risk players missing out on your AI systems, what other risks will you mitigate?

References

- A Sharp, LLC. 1999. *King of Dragon Pass*. A Sharp, LLC.
A Sharp, LLC. 2017. *Six Ages*. A Sharp, LLC.

- Adams, T., and Adams, Z. 2006. *Slaves to Armok: God of Blood Chapter II: Dwarf Fortress*. Bay 12 Games.
- Adams, T. 2017. Secret identities in Dwarf Fortress. In *Proc. Experimental AI in Games 4*.
- Digman, J. M. 1990. Personality structure: Emergence of the five-factor model. *Annual Review of Psychology* 41(1):417–440.
- Dynamic Pixels. 2017. *Hello Neighbor*. tinyBuild Games.
- Epictetus, and Higginson, T. W. 1944. *Epictetus: Discourses and Enchiridion, Based on the Translation of Thomas Wentworth Higginson*. Walter J. Black, Inc.
- Firaxis Games. 2016. *Civilization VI*. 2K Games.
- Funcom. 2008. *Age of Conan: Hyborian Adventures*. Funcom.
- Kitfox Games. 2016. *Moon Hunters*. Kitfox Games.
- Kitfox Games. 2017. *The Shrouded Isle*. Kitfox Games.
- Maslow, A. H. 1943. A theory of human motivation. *Psychological Review* 50(4):370.
- Maxis. 2014. *The Sims 4*. Electronic Arts.
- Paradox Development Studio. 2012. *Crusader Kings II*. Paradox Interactive.
- Roberts, B. W.; Walton, K. E.; and Viechtbauer, W. 2006. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin* 132(1):1.
- The Tiniest Shark. 2013. *Redshirt*. Positech.