Why Non-Specific Simulation Works

The research behind GemaSim

GemaSim NZ Ltd

April 2026

How this paper came about

This paper began as an investigation rather than an argument. The work behind GemaSim has always been grounded in cognitive science and adult learning theory, but those grounds had not been laid out in a single place where they could be examined. The simulation field has also moved over the last decade, and some of the assumptions that shaped early thinking about leadership simulation needed to be retested against the current evidence.

Five questions were put to the literature. Each was treated as a real question, not a rhetorical one. The findings below are what the evidence supports, what it suggests, and where it does not yet establish what we might wish.

1. Does training transfer across stripped contexts?

Does practice in a deliberately abstract setting produce skills that travel to real workplaces, or does context-stripping produce inert knowledge that stays where it was learned?

2. Does the high-fidelity argument apply to non-technical skills?

The dominant assumption in much of the simulation industry is that more realism produces more transfer. The question is whether this is true for the behavioural skills, teamwork, communication, decision-making under stress, that leadership simulation is usually trying to develop.

3. What kinds of skills are actually trainable this way?

Not all skills behave the same. The question is which categories of skill are well served by non-specific simulation, and which are not.

4. Is the debrief the active ingredient?

A long-standing tension in the simulation literature is whether the simulator itself or the structured conversation afterwards does the work. The question is what the evidence actually says.

5. Does practising under pressure build performance under pressure?

Behavioural patterns are most visible under pressure, and the simulation is designed to make them visible by producing real load. The question is whether practising non-technical skills under pressure prepares people to use those skills when the pressure is real, or whether the experience stays in the simulator.

The short answers

The longer answers are in the chapters that follow.

Transfer across stripped contexts is supported by a deep and consistent cognitive science of analogical learning, with a narrower applied evidence base specifically in working professionals.

The high-fidelity assumption is empirically weak for non-technical skills, and the field's leading methodologists have argued for retiring the term fidelity altogether.

Generic teamwork, stress regulation and behavioural pattern awareness are well supported as trainable through this design. Domain-specific decision expertise is not, and requires in-context experience that no generic simulation can substitute.

The structured debrief is the most consistently identified active ingredient across multiple major meta-analyses.

Practising under pressure builds performance under pressure. The evidence is direct: a meta-analysis of pressure training across sport and law enforcement, alongside experimental work showing that skills practised under anxiety hold up under anxiety while the same skills practised in calm do not.

What follows works through each of these in turn.

Skills that don't travel

A team finishes a leadership programme. Glowing scores on the feedback forms. Three months later, the same patterns are back. The senior manager who was supposed to delegate more is back to firefighting. The team that practised difficult conversations is avoiding the same difficult conversations. The training happened. The change did not.

This is a transfer problem. People learn things in one context and cannot move them to another. The cognitive science literature has been describing this for decades, and the pattern is robust enough to have its own taxonomy. Barnett and Ceci (2002) catalogued the dimensions across which transfer is hard: the knowledge domain, the physical setting, the social context, the temporal gap, the modality of the task. Move the learning across enough of these at once and most of it stays behind.

The dominant response in corporate training has been to make the learning context look more like the application context. Industry-specific scenarios. Role-plays drawn from the sector. Case studies of companies just like the client's. The intuition is reasonable: if the practice looks like the real thing, it should transfer to the real thing. The intuition is also wrong, in a particular and well-documented way.

The research suggests something almost backwards. Stripping the surface context away can support better transfer than matching it. Not because realism is unimportant, but because what is usually meant by realism (the props, the industry vocabulary, the recognisable setting) is not what makes learning travel. What makes learning travel is whether the underlying pattern was seen clearly enough to be applied somewhere new. And surface context can actively get in the way of that.

Which skills are we talking about

Not all skills behave the same way. A surgical procedure transfers very differently from the ability to stay calm under pressure. Both can be trained, but the design that works for one will not work for the other.

Two kinds of expertise have been distinguished in the literature for nearly forty years (Hatano and Inagaki, 1986). Routine expertise is high speed and high accuracy on familiar tasks in familiar conditions. Adaptive expertise is the ability to apply understanding flexibly when conditions change. The two need different training. Routine expertise grows from extensive practice in the actual context of use. Adaptive expertise grows from seeing the same underlying pattern in many different surface forms.

There is also a long-running scholarly tradition, naturalistic decision making (Klein, 2008), which is clear about its boundary: deep, domain-specific decision expertise is built from extensive in-context experience. A pilot recognising a cabin pressure problem from a thousand small cues. A paediatrician spotting a sick infant before any single sign confirms it. That kind of pattern recognition cannot be transferred from a generic training session. It is grown in the actual cockpit, the actual ward.

What can be trained across contexts is a different layer. Generic teamwork. Stress regulation. Behavioural pattern awareness. The ability to notice one's own and others' shifts under pressure. The ability to communicate clearly when the situation is moving fast. These are the non-technical skills that high-reliability industries have been formalising for decades, with converging taxonomies across aviation, surgery, anaesthesia and nuclear operations (Flin, O'Connor and Crichton, 2008). They are the layer that transfers, and they are the layer that non-specific simulation is designed to address.

How abstraction makes learning travel

The original work on this is from the 1980s. Gick and Holyoak (1983) gave students a problem about destroying a tumour with radiation, dressed up in the surface details of cancer treatment. Most students could not solve it. The same students were then given an analogous story first, about a general attacking a fortress, before the medical problem. Now they could solve it. The story and the problem had nothing in common on the surface, but the underlying structure was the same. Once both were in view, the structure became visible.

What unlocked the transfer was not similarity to the target problem. It was comparison across structurally similar but surface-different examples. Catrambone and Holyoak (1989) followed up and showed that the comparison had to be active. Studying two cases sequentially did not produce transfer. Comparing them, with attention drawn to the underlying principle, did. The mechanism is called schema abstraction: when the same pattern shows up in different clothes, the pattern itself becomes the thing the learner remembers.

This finding has held up across forty years of replications. It works in mathematics classrooms (Richland, Zur and Holyoak, 2007). It works with adult professionals learning negotiation (Gentner, Loewenstein and Thompson, 2003). It works in virtual training environments where it can be measured directly: in one recent study, high-abstraction training produced better generalisation to new tasks than low-abstraction training, with EEG showing higher initial cognitive load that decreased as the abstraction was internalised (Hernández-Castañón et al., 2023). The cost of abstract training shows up early. The benefit shows up later, when the situation changes.

A reasonable question follows. If structural variation across surface contexts is what supports transfer, what about a simulation system that uses the same wrapper every time? GemaSim places every team on a series of space missions. The cockpit, the role of mission commander, the oxygen and points and screens, those things are constant. So is the abstraction argument intact?

The argument depends on understanding what is being abstracted from what. The thing GemaSim is trying to make visible is not how to fly a spaceship. The space wrapper is the source of analogy, not the subject of training. What varies across missions is the behavioural challenge: who holds authority, who depends on whom, what time pressure looks like, what kind of breakdown happens, what kind of decision is needed, who can see what. Each mission gives the brain a different instance of the same underlying behavioural pattern. The pattern is what gets abstracted. The space context is the constant frame that lets the variation become visible.

There is a second function the constant wrapper serves. It is uniformly distant from every learner's actual workplace. Nobody works on a spaceship. Which means nobody can dismiss what they see in the simulation as a feature of someone else's industry. A finance team and a hospital team flying the same mission see the same patterns surface in themselves and in each other. The neutral ground is what makes the comparison honest.

The fidelity question

If high-fidelity replication is so weakly linked to transfer of non-technical skills, why is it the dominant assumption in the simulation industry? Partly because fidelity is visible. A high-fidelity manikin or a fully-rigged operating room is a tangible product that can be specified, photographed, and procured. A trainer's debrief skill cannot be. What is visible tends to drive what gets built and bought.

The empirical position has been moving, though, and the people who built the major simulation meta-analyses have been at the centre of the move. Cook and colleagues (2011) reviewed 609 studies covering 35,226 trainees. They confirmed that simulation as a category produces large effects on knowledge, skills and behaviour. They also found that subgroup analyses of specific design features did not produce consistent significant effects. Whatever simulation is doing, it is not particularly sensitive to which version of high fidelity is chosen.

Three years later, several of the same authors went further. In a perspective titled Reconsidering Fidelity in Simulation-Based Training (Hamstra et al., 2014), they recommended dropping the term fidelity altogether. Their argument was not just that the empirical relationship is weak. It was that the term itself confuses physical resemblance with functional

task demand, and that this confusion is built into how the field talks about its own work. They proposed reframing around transfer and task demand instead.

The pattern shows up in the specific evidence on non-technical skills. Norman, Dore and Grierson (2012) reviewed the medical education literature and concluded that conceptual fidelity, whether the cognitive demands match the target task, matters more than physical fidelity, whether the simulator looks like the real thing. Gu and colleagues (2017) ran a randomised non-inferiority trial: thirty-six anaesthesiology trainees who trained on a low-fidelity simulator acquired non-technical skills as effectively as those who trained on a high-fidelity simulator. A meta-analysis of 57 troubleshooting training experiments (Doozandeh and Hedayati, 2022) found no single fidelity level was universally superior; the right level depended on prior skill and task type.

Even the professional standards have moved. The 2021 INACSL Healthcare Simulation Standards of Best Practice, the most widely adopted simulation standards in healthcare, recognise three dimensions of fidelity (physical, conceptual, psychological) and frame them as working together rather than as a hierarchy. The standards explicitly state that simulation design "is less about specific reality and should instead focus on representing stimuli and cues that would typically be present to drive decision-making and action." The intuition that more realism is always better is not where the field has landed.

The debrief is what works

Experiential learning, broadly, outperforms passive instruction. People remember more, apply more, and change behaviour more after doing something than after being told about it. This is not in serious dispute. The interesting question is what makes one experiential design produce transfer and another produce a memorable afternoon. The most consistent answer in the literature is also the most unfashionable: it is the structured conversation afterwards.

Tannenbaum and Cerasoli (2013) ran a meta-analysis across military, medical, aviation and corporate settings: 46 samples, 2,136 participants. Structured debriefs produced a moderate-to-large performance improvement (d=0.67) over experiential learning without debrief. The effect was robust across very different industries and very different debrief formats. Whatever variation exists between specific debrief methods, the presence of a structured debrief is what consistently shows up as the difference-maker.

Cheng and colleagues (2014) confirmed and extended the picture in healthcare. Their meta-analysis covered 177 studies and 11,511 learners. Effect sizes for simulation-with-debrief versus no intervention ranged from 0.28 to 2.16 across outcomes, all favouring debriefed

simulation. When they compared specific design features against each other (video-assisted versus not, long versus short), most differences were small or non-significant. The finding is not that one debrief style dominates the others. The finding is that having a structured debrief at all is what matters.

Sawyer and colleagues (2016) reviewed the established debrief methods (facilitator-guided approaches like Debriefing with Good Judgement, structured frameworks like PEARLS, learner-driven formats) and reached a similar conclusion. Multiple methods have evidence behind them. None has been established as universally superior. The match should be to the learning goals and the experience level of both learners and facilitators.

Wouters et al. (2013) provides a useful negative case. Their meta-analysis of 77 serious games found a moderate overall effect (d=0.29 for learning, 0.36 for retention) which got stronger when games were combined with other instruction, used over multiple sessions, and included group work. A serious game in isolation, used once, with no debrief, is not the design the evidence supports. The simulation is the experience. The debrief is what turns experience into learning that travels.

This is also where the contested claims about decontextualised simulation become defensible. The strong-form situated cognition view (Brown, Collins and Duguid, 1989) argues that knowledge stripped from its context becomes inert. The objection is real, and it would be fatal to a model that left the learner alone in a generic simulation and expected transfer to materialise. What makes non-specific simulation work is that it does not stop at the experience. The structured debrief explicitly bridges the abstract experience back to the learner's actual context. The decontextualised mission becomes the source case; the debrief is where the analogical comparison happens; the learner's own workplace becomes the target. This is the contextualise-decontextualise-recontextualise cycle that Abraham and colleagues (2024) propose as the design pattern for non-technical skills training.

Stress, and what transfers

Behavioural patterns are most visible, and most consequential, under pressure. Calm-water leadership and white-water leadership are different things, and the second is what people are usually being trained to handle. The relevant question is whether practising non-technical skills under pressure prepares people to use those skills when the pressure is real.

On this, the evidence is direct. Low and colleagues (2021) meta-analysed pressure training across sport and law enforcement: fourteen studies, thirteen of which showed that participants who trained under pressure outperformed controls when subsequently tested under pressure.

The effect (Hedges' g = 0.67) is moderate to large, and it held across both novices and experts. The earlier experimental work that the meta-analysis builds on is just as clear. Oudejans and Pijpers (2009) trained expert basketball players and expert dart players either with anxiety or without it, then tested both groups under high anxiety. The group that trained without anxiety choked. The group that trained with anxiety did not. Lawrence and colleagues (2014) showed the same effect with a larger sample and named the principle: skills practised under anxiety transfer to anxious performance better than the same skills practised in calm.

The mechanism has two well-established components. Repeated exposure to a stressor produces habituation: the hypothalamic-pituitary-adrenal axis response to a recurring stressor declines with repetition (Grissom and Bhatnagar, 2009). What people experience as getting used to it is real and physiological. The person who has been through twenty high-stakes situations responds with less cortisol, less heart-rate spike, less attentional narrowing than they did the first time. Alongside this, skills practised under load become more robust to load. The cognitive and motor systems learn to function with the additional demand of stress already factored in, rather than collapsing when the stress arrives unexpectedly.

What this matters for is which patterns the practice has the chance to address. Under arousal, attention narrows and peripheral cues are missed (Easterbrook, 1959). In a team context, the narrowing extends to the team itself: stressed team members shift from a broad team perspective to an individualistic self-focus, and this loss of team perspective mediates the decline in team performance (Driskell, Salas and Johnston, 1999). These are the patterns that surface in the simulation, and they are the same patterns that surface at work. The simulation that produces real cognitive and emotional load surfaces them in a place where they can be seen and discussed; the simulation that produces only theatre does not.

What the evidence requires of the simulation is real load. The pressure does not have to be extreme; mild anxiety in training transfers to performance under higher anxiety (Oudejans and Pijpers, 2010). The pressure does have to be genuine. If the cognitive and emotional load is real, the stress response habituates, the skills become more automated under load, and the patterns that show up in the simulation are the patterns that show up at work. If the load is artificial, none of those things happens.

Where this fits in a learning system

Non-specific simulation is not a complete training programme on its own. It is one component, with a specific job, sitting alongside others that do different jobs.

What it is good for: surfacing behavioural patterns under stress, training generic teamwork and coordination, building stress regulation, developing metacognitive awareness of one's own patterns. The evidence supports each of these. Generic teamwork training produces team-level cognitive and skill gains (Ellis et al., 2005). Healthcare team training meta-analyses show transfer all the way through to organisational outcomes (Hughes et al., 2016; Salas et al., 2008). Reviews of teamwork in healthcare confirm that generic competencies are trainable across clinical teams while domain-specific decision expertise still requires domain experience (Rosen et al., 2018).

What it is not good for: building deep domain knowledge, training procedural skills, replacing in-context experience for naturalistic decision making. A pilot still needs to fly aircraft. A surgeon still needs to operate. A new manager still needs to manage real teams. Non-specific simulation accelerates the behavioural layer underneath those activities; it does not replace the activities themselves.

This is the scope of the claim. A particular kind of training, well designed and well facilitated, helps people see and modify their own behavioural patterns under pressure in a way that transfers to other pressured settings. It does not change a career in one session, and it does not substitute for the experience of doing the real job.

What the evidence does and does not show

Strong evidence: the cognitive mechanism of analogical transfer through structural comparison has held up across forty years of replication. The role of structured debriefing as the active ingredient in simulation learning is established by multiple major meta-analyses across industries. Team training transfers to organisational outcomes. The relationship between high physical fidelity and non-technical skill transfer is empirically weak; the field's leading methodologists have argued for retiring the term fidelity altogether. Practising non-technical skills under pressure builds the capacity to use them under pressure, with meta-analytic support from sport and law enforcement and converging experimental evidence on the specificity of practice and on stress habituation.

Suggestive evidence: the specific claim that decontextualised simulation outperforms contextualised simulation in working professionals rests primarily on a single empirical study (Fauquet-Alekhine and Boucherand, 2016, N=33) and the conceptual framework that interprets it (Abraham et al., 2024). The applied evidence base is narrower than the analogical-transfer foundation it draws on. The position is best defended on the strength of the underlying mechanism, not on the volume of direct empirical replication.

Important limits: the evidence does not support a claim that non-specific simulation builds domain-specific decision expertise (Klein, 2008). That requires immersion in the actual domain. What this kind of simulation trains is the generic, transferable layer: stress regulation, behavioural pattern awareness, and team coordination skills. Strong-form situated cognition (Brown, Collins and Duguid, 1989) raises a real challenge for any decontextualised approach. The honest response is that effective non-specific simulation is not pure decontextualised instruction; the structured debrief is what addresses the situated cognition concern by recontextualising abstract experience back into the learner's real workplace.

What this looks like when it works

The case for non-specific simulation is structural. It explains why a particular design can produce learning that travels. It does not, by itself, prove that any given session will. The proof of that lives in the room.

What it looks like when it works is something most of the people who run these sessions can describe without hesitation. A team that has just finished its third mission, sitting in a debrief, recognising patterns in itself that nobody had named before. The senior person noticing that they took authority back the moment the situation got tense, and that they do this at work too. The quiet team member realising that their silence under pressure is read as agreement, and that they have been agreeing to things they did not agree with for years. The two colleagues who realise they have been working past each other on every difficult decision, and seeing why for the first time.

These are not climactic moments. They are small, dry, often awkward. But they are the moments where behavioural patterns become visible to the people inside them, and where the possibility of a different pattern becomes real. Cognitive science explains why this can happen in a stripped-down space simulation as readily as in the team's own meeting room, and often more readily. The mechanism is the same one that has been studied for forty years. The evidence is what the evidence is.

What the evidence cannot do is replace the trainer's craft, the team's willingness to look honestly at what surfaces, or the work of taking what was seen in the simulation back into the workplace where the real patterns live. The science explains why the design can work. Whether it does work, in any particular room, depends on the people in the room.

GemaSim is one implementation of the design. The case for the approach does not rest on GemaSim itself. It rests on whether the underlying argument about how transferable skills are learned holds up. The argument does. The rest is craft.

References

Abraham, P., et al. (2024). Enhancing perioperative care through decontextualized simulation: A game-changer for non-technical skills training. Journal of Clinical Anesthesia, 94, 111428.

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. Educational Researcher, 25(4), 5–11.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. Psychological Bulletin, 128(4), 612–637.

Beaubien, J. M., & Baker, D. P. (2004). The use of simulation for training teamwork skills in health care: How low can you go? Quality and Safety in Health Care, 13(Suppl 1), i51–i56.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 32–42.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(6), 1147–1156.

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D. A. (2014). Debriefing for technology-enhanced simulation: A systematic review and meta-analysis. Medical Education, 48(7), 657–666.

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., Erwin, P. J., & Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. JAMA, 306(9), 978–988.

Doozandeh, P., & Hedayati, S. (2022). The effect of simulation fidelity on transfer of training for troubleshooting professionals: A meta-analysis. IISE Transactions on Occupational Ergonomics and Human Factors, 10(3).

Driskell, J. E., Salas, E., & Johnston, J. (1999). Does stress lead to a loss of team perspective? Group Dynamics: Theory, Research, and Practice, 3(4), 291–302.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. Psychological Review, 66(3), 183–201.

Ellis, A. P. J., Bell, B. S., Ployhart, R. E., Hollenbeck, J. R., & Ilgen, D. R. (2005). An evaluation of generic teamwork skills training with action teams: Effects on cognitive and skill-based outcomes. Personnel Psychology, 58(3), 641–672.

Fauquet-Alekhine, P., & Boucherand, A. (2016). Decontextualization effect in simulation training for transverse professional practices. British Journal of Education, Society and Behavioural Science, 14(2), 1–17.

Flin, R., O'Connor, P., & Crichton, M. (2008). Safety at the sharp end: A guide to non-technical skills. Ashgate.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. Journal of Educational Psychology, 95(2), 393–408.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15(1), 1–38.

Grissom, N., & Bhatnagar, S. (2009). Habituation to repeated stress: Get used to it. Neurobiology of Learning and Memory, 92(2), 215–224.

Gu, Y., Witter, T., Livingston, P., Rao, P., Varshney, T., Kuca, T., & Bould, M. D. (2017). Anaesthesiology residents' performance and the effect of simulator fidelity on acquiring non-technical skills: A randomized non-inferiority trial. Canadian Journal of Anesthesia, 64(12), 1182–1193.

Hamstra, S. J., Brydges, R., Hatala, R., Zendejas, B., & Cook, D. A. (2014). Reconsidering fidelity in simulation-based training. Academic Medicine, 89(3), 387–392.

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), Child development and education in Japan. W. H. Freeman.

Hernández-Castañón, V. del R., et al. (2023). Effect of the level of task abstraction on the transfer of knowledge from virtual environments in cognitive and motor tasks. Frontiers in Behavioral Neuroscience, 17, 1162744.

Hughes, A. M., Gregory, M. E., Joseph, D. L., et al. (2016). Saving lives: A meta-analysis of team training in healthcare. Journal of Applied Psychology, 101(9), 1266–1304.

INACSL Standards Committee. (2021). Healthcare simulation standards of best practice: Simulation design. Clinical Simulation in Nursing, 58, 14–21.

Klein, G. (2008). Naturalistic decision making. Human Factors, 50(3), 456–460.

Lawrence, G. P., Cassell, V. E., Beattie, S., Woodman, T., Khan, M. A., Hardy, L., & Gottwald, V. M. (2014). Practice with anxiety improves performance, but only when anxious: Evidence for the specificity of practice hypothesis. Psychological Research, 78(5), 634–650.

Low, W. R., Sandercock, G. R. H., Freeman, P., Winter, M. E., Butt, J., & Maynard, I. (2021). Pressure training for performance domains: A meta-analysis. Sport, Exercise, and Performance Psychology, 10(1), 149–163.

Norman, G., Dore, K., & Grierson, L. (2012). The minimal relationship between simulation fidelity and transfer of learning. Medical Education, 46(7), 636–647.

Oudejans, R. R. D., & Pijpers, J. R. (2009). Training with anxiety has a positive effect on expert perceptual-motor performance under pressure. Quarterly Journal of Experimental Psychology, 62(8), 1631–1647.

Oudejans, R. R. D., & Pijpers, J. R. (2010). Training with mild anxiety may prevent choking under higher levels of anxiety. Psychology of Sport and Exercise, 11(1), 44–50.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. Science, 316(5828), 1128–1129.

Rosen, M. A., DiazGranados, D., Dietz, A. S., et al. (2018). Teamwork in healthcare: Key discoveries enabling safer, high-quality care. American Psychologist, 73(4), 433–450.

Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does team training improve team performance? A meta-analysis. Human Factors, 50(6), 903–933.

Sawyer, T., Eppich, W., Brett-Fleegler, M., Grant, V., & Cheng, A. (2016). More than one way to debrief: A critical review of healthcare simulation debriefing methods. Simulation in Healthcare, 11(3), 209–217.

Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. Human Factors, 55(1), 231–245.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. Journal of Educational Psychology, 105(2), 249–265.