RESEARCH PAPER

Wit and wisdom: using computational humor to communicate about economics

Iacob Postavaru[1], Emilia Bunea[2], Crina Pungulescu[3] and David Stolin[4]

[1] Amsterdam University College, Science Park 113, Amsterdam 1098 XG
[2] Vrije Universiteit Amsterdam, De Boelelaan 1105, Amsterdam 1081 HV
[3] John Cabot University, Via della Lungara 233, 00165 Rome
[4] TBS Business School, 1 Place Alphonse Jourdain - CS 66810, 31068 Toulouse

Submission date: 18 March 2024; Acceptance date: 20 May 2024; Publication date: 6 March 2024

ABSTRACT

This paper explores the potential of large language models to enhance economics education through computational humor. We employ OpenAI's GPT-4 model to infuse humor into summaries of three Nobel laureates' contributions to economics and conduct a small empirical exercise with undergraduate students to test the pedagogical efficacy of computational humor. The results suggest that computer-generated humor may be an effective learning aid: the results of the students who rate the humorous versions of the instructional texts as genuinely funny are significantly better than the results of their peers who are not amused. Encouragingly for teachers who try to be funny but fail, we do not find evidence that ineffectual humor is detrimental to learning.

KEYWORDS

natural language processing, large language models, instructional humor

Introduction

Large language models already perform successfully in supportive roles in Economics research, for instance when tasked with proposing titles and writing abstracts for academic papers (Ash and Hansen, 2023). As they continue to improve, we must face the intriguing possibility of casting them in leading roles. In this paper, OpenAI's GPT-4 model[1] auditions for such a role, as it is asked to enhance standard instructional texts in economics through judicious use of humor.

This paper investigates whether entirely autonomously generated computational humor can impact learning experience. (Very) large language models have undergone enormous

[1]https://openai.com/gpt-4

CONTACT: d.stolin@tbs-education.fr
ACCEPTING EDITOR: Peter Senker

improvements recently and are able to perform high-level actions with such accuracy that GPT-3 has earned an IQ of 150 (Ray, 2023). Humor, which is infinitely versatile by nature, seems an adequate test, heightened by the challenges of an educational context.

We generated the pipeline that prompts the model to produce adequate instructional humor to enhance three instructional texts in economics. The instructional texts summarize the contributions of three Nobel prize-winning theorists: Oliver Hart, Bengt Holmström and Paul Milgrom. The reason for this choice is that all three had accepted to participate in a comedian-presented (and therefore more light-hearted than usual) panel discussion which the authors of the study helped organize at the bicentennial meeting of the European Finance Association in August 2023 in Amsterdam.[2] By focusing on these particular laureates, we test the effect of seeing short clips of the laureates in a humorous context on the understanding of their work.[3]

We tested the pedagogical efficacy of the GPT-4 generated model in a sample of 52 undergraduate students and found that simply exposing students to the version of the instructional texts enhanced by computational humor is not sufficient to induce a significant difference in performance. What matters is that the attempt at humor be successful: only the students who find the 'humorous' text actually 'funny' have significantly better results.[4] However, these encouraging findings are driven by the sense of humor and performance of a subset of only seven out of the 26 respondents who were exposed to the humorous version of the instructional texts. Thus, we acknowledge the small sample limitations of our empirical exercise, and welcome further scrutiny of this new and potentially very rewarding field of research, which enlists the high-level capabilities of large language models in the service of economics education.

Theoretical considerations

Humor in education

Numerous studies have attempted to determine whether humor enhances learning, with mixed results. On the one hand, instructional humor processing theory (IHPT) proposes that instructor humor increases recall and learning, provided that (a) the humor is relevant to the instructional contents; (b) students actually perceive it as funny; and (c) students are motivated and able to process the instructional message (Wanzer *et al.*, 2010). On the other hand, several studies have exposed students to either standard or humorous examples and found that the students in the latter group perform worse. For example, Bolkan *et al.* (2018) conclude that, when humor is integrated in the instructional lessons, it competes for student learning with the concepts being taught. In their view, contiguous humor (not linked directly to the content) may provide experiential and motivational benefits with less risk for learning outcomes. However, Bolkan *et al.* (2018) do not test whether students who perceived the humorous examples as funny performed worse or better than those who did not or than those in the control group.

Given the importance of the content and quality of the humor involved, it is surprising how little of the research on instructional humor is based on humor generated by the authors (rather than on polling students about their experience in the classroom). Of the handful of studies that do generate their own humorous material, most are vague on how this was done (e.g., Celik and Gungdogdu, 2016; Buttussi and Chittaro, 2020; Erdogdu and Cakiroglu, 2021), which not only impedes replicability, but offers little practical help to instructors.

---

[2]https://efa2023.efa-meetings.org/

[3]We note that the full panel discussion has since been made publicly available on https://youtu.be/5o9jGgmDwos

[4]It is also possible that unobserved student characteristics are associated with better understanding as well as with greater enjoyment of the humor.

Hypotheses

Given the challenges of generating adequate and effective humor for the classroom and the replicability issues already mentioned, this paper opens the way towards investigating whether and how large language models can help. While artificial intelligence is already an established feature in higher education for such tasks as automatic question generation and grading, as well as intelligent tutoring systems designed to provide individualized feedback to students (see Crompton and Burke, 2023, for a review), to our knowledge, its ability to produce effective instructional humor has not yet been put to the test. We hypothesize that:

- large language models produce adequate instructional humor, and that
- instructional humor enhances learning.

We test these hypotheses in a pilot experiment.

Empirical analysis

Computational humor

Large language models can be leveraged to act as agents following predefined objectives. In this study, we use OpenAI's GPT-4 model to create an entirely autonomous method for generating instructional humor. The model has been tasked to integrate humor into the content through a metaphor, anecdote or quip, while at the same time avoiding humor-induced ambiguity. To mitigate potential errors while preserving the autonomy of the process, we draw on the methodological contributions of Shinn *et al.* (2023) and Nair *et al.* (2023) and add an iterative 'self-reflection' feature to the pipeline: the model will self-evaluate and improve through several iterations. The original (input) text passes through three transformative steps before the final output is produced:

- generator – enhances the text with instructional humor
- evaluator – lists the pros and cons of the generated text, and
- decider – selects the best option.

The generator and evaluator functions have been assigned different roles: university professor in economics with experience and achievements in humorous teaching, researcher in educational humor and comedian with a background in economics. The decider function plays the role of a professor experienced in the evaluation of educational humor. To illustrate the validity of the process, we provide the following excerpt from the input text:

> In his analysis on how a CEO's contract should be formulated, Holmström proposed the 'multi-tasking model', which acknowledges the complexity of a CEO's role and the various tasks they need to perform.

This excerpt, enhanced by computational humor, becomes:

> When it comes to our multitasking maestros, the CEOs, Holmström came up with the aptly named 'multi-tasking model'. This model acknowledges that a CEO's role is as complex as a Rubik's cube, with various tasks that need to be tackled effectively.

We found that GPT-4 can be capable of producing adequate instructional humor, in line with our first hypothesis. Although only seven out of our 26 respondents who received the humorous versions of the instructional texts rated the GPT-4 humor as actually funny, they are also the ones whose performance was better. Given the vast differences in humor appreciation across individuals (see

Warren *et al.*, 2020, for a comprehensive review), we make all the instructional texts available[5] so that readers can judge GPT-4's comic aptitude for themselves.

Experimental design

Some 52 respondents aged 18 to 35 and currently enrolled in undergraduate programs were recruited through the Prolific survey platform.[6] The participants were screened in order to achieve equal gender distribution and then randomly assigned to one of two groups: the control group, which received three original instructional texts (input to the GPT-4 pipeline) and the test group, which received the GPT-4 output texts, enhanced by computational humor. The texts (approximately 800 words long) were followed by a quiz consisting of 17 multiple choice questions and a short survey designed to elicit the respondents' perceptions (on a five-point scale), along several dimensions:

- familiarity with economics – (1) 'not at all familiar' to (5) 'extremely familiar'
- attention – (1) 'did not capture my attention' to (5) 'did capture my attention'
- excitement – (1) 'boring' to (5) 'exciting'
- interest – (1) 'not at all interesting' to (5) 'very interesting'
- humor – (1) 'serious' to (5) 'humorous'
- fun: (1) 'not funny' to (5) 'funny'.

Results

Respondents who found the texts 'interesting' tend to get better quiz results, with the pairwise correlation coefficient between the two variables of 0.65 (see Table 1). Familiarity with economics and the ability of the text to capture the reader's attention also correlate positively with the quiz results (yet moderately, with correlation coefficients of 0.36 and 0.34, respectively).

Table 1 reports the pairwise correlations for our variables of interest – quiz results, familiarity with economics, attention, excitement, interest, humor and fun – for the full sample of 52 respondents. Statistical significance is denoted by *** (at 1%), ** (at 5%) and * (at 10%).

Table 1. Pairwise correlations – full sample

	Quiz results	**Familiarity econ.**	**Attention**	**Excitement**	**Interest**	**Humor**	**Fun**
Quiz results	1	0.34**	0.36***	0.29**	0.65***	−0.05	0.09
Familiarity econ.		1	0.44***	0.47***	0.50***	0.15	0.27*
Attention			1	0.75***	0.36***	0.13	0.10
Excitement				1	0.65***	0.22	0.13
Interest					1	0.14	0.01
Humor						1	0.65***
Fun							1

A text that is perceived as 'exciting' appears to be more successful in capturing attention (with a full sample correlation coefficient of 0.75 between the two variables) than a text that is labeled 'interesting' (where correlation with attention is only 0.36). This difference suggests that the personal (potentially more emotionally charged) endorsement of 'exciting' carries more weight

[5]Appendices A, B and C present the full texts received by the participants in the control and the test group respectively.
[6]https://www.prolific.com/

than the more detached (intellectual) characterization of 'interesting'. A similar distinction in terms of the respondents' personal experience may be inferred from the assignment of a 'funny' versus 'humorous' label to the text (which, just like the 'excitement' and 'interest' variables correlate strongly, but not overwhelmingly, at 0.65). An unsuccessful attempt to amuse may still be recognized as 'humorous' even when it falls short of 'funny', as suggested by the fact that the average ratings are higher for the 'humor' variable than for the 'fun' variable within each of the groups (see Table 2). In the same vein, five of the 26 respondents in the control group (who received the input text) gave a high rating (3 or above) for the 'humorous' variable, while, unsurprisingly, none found it 'funny' (all the ratings for the 'fun' variable are either 1 or 2) (see Table 3). Notably, we do not find any evidence of a detrimental effect of instructional humor on learning, as the difference between the average quiz results of the control and test group (78.51 and 75.79) is very small and highly insignificant (with a *p*-value of 0.60).

Table 2 reports the number of observations and average values for our variables of interest – quiz results, familiarity with economics, attention, excitement, interest, humor and fun – for the full sample as well as for the control and test groups. *P*-values for tests of significance for the difference in the means of the variables for the control versus test groups are reported both for the standard *t*-test (assuming equal population variances) and the Satterthwaite-Welch *t*-test, which allows unequal variances in the two populations. Statistical significance is denoted by *** (at 1%), ** (at 5%) and * (at 10%).

Table 2. Descriptive statistics

	Full sample	**Control group**	**Test group**	***t*-test**	***t*^{SW}-test**
No. obs.	52	26	26	*p*-value	*p*-value
Quiz results	77.15	78.51	75.79	0.6076	0.6077
Familiarity econ.	2.25	2.04	2.46	0.1536	0.1540
Attention	3.12	3.12	3.12	1.0000	1.0000
Excitement	2.58	2.62	2.54	0.7989	0.7989
Interest	3.33	3.54	3.12	0.1807	0.1807
Humor	1.98	1.73	2.23	0.1110	0.1112
Fun	1.56	1.15	1.96	0.0005***	0.0007***

Table 2 shows that the average grade for 'fun' is the lowest of all the variables measured, at only 1.96. Yet it seems there is potential in successfully striking the 'fun' chord, as the average quiz results of the seven respondents who gave high marks (3 or above) for the 'fun' component are the highest of all (87.39), slightly exceeding the results of the respondents who were very familiar with economics (86.93 in the control group and 86.63 in the test group). It is plausible that familiarity with economics plays a role in humor appreciation as well as in performing well in the quiz. However, we note that the seven respondents who gave GPT-4 high marks for humor have on average lower familiarity with economics (3.14) and slightly higher performance than the respondents in the test group who are familiar with economics (3.63). Given the size of the sample, results should be interpreted with caution, but the fact that the respondents from the test group who declare themselves not amused get significantly lower quiz results, with an average of 71.52 (see Table 4), is encouraging for our second hypothesis: humor appears to make a difference in learning only when it is perceived as actually funny. The average result of the students who gave low marks to the 'fun' component – just like the average result for the respondents in the same group who gave low ratings for the 'humor' content (73.01) – is comparable (more often than not, favorably) with the average results obtained in the 'low' sections for 'familiarity with economics', 'attention', 'excitement' and 'interest' of both the control and the test groups.

Table 3 reports the number of respondents that give high (3 or above) versus low (1 or 2) ratings to the following variables – familiarity with economics, attention, excitement, interest, humor and fun – for the full sample as well as for the control and test groups.

Table 3. High vs. low ratings subgroups

	Full sample		**Control group**		**Test group**	
	High	**Low**	**High**	**Low**	**High**	**Low**
Familiarity econ.	20	32	9	17	11	15
Attention	36	16	18	8	18	8
Excitement	26	26	11	16	15	11
Interest	41	11	22	4	19	7
Humor	14	38	5	21	9	17
Fun	7	45	0	26	7	19

Table 4 reports average quiz results for the subgroups that give high (3 or above) and low (1 or 2) ratings to the following variables – familiarity with economics, attention, excitement, interest, humor and fun for the full sample as well as for the control and test groups. *P*-values for tests of significance for the difference in the means of the variables for the control versus test groups are reported for both the standard *t*-test (assuming equal population variances) and the Satterthwaite-Welch *t*-test, which allows unequal variances in the two populations. Statistical significance is denoted by *** (at 1%), ** (at 5%) and * (at 10%).

Table 4. Quiz results

	Full sample				**Control group**				**Test group**			
	High	**Low**	**t-test**	**t^{SW}-test**	**High**	**Low**	**t-test**	**t^{SW}-test**	**High**	**Low**	**t-test**	**t^{SW}-test**
Familiarity econ.	86.76	71.14	0.0027***	0.0006***	86.93	74.05	0.1235	0.0520*	86.63	67.84	0.0047***	0.0027***
Attention	82.19	65.81	0.0028***	0.0130**	82.68	69.12	0.1152	0.1729	81.70	62.50	0.0074***	0.0425**
Excitement	81.45	72.85	0.0997*	0.0999*	82.35	75.69	0.4161	0.4254	80.78	68.98	0.0921*	0.1242
Interest	79.77	67.38	0.0514*	0.1205	80.21	69.12	0.3213	0.5351	79.26	66.39	0.0995*	0.1597
Humor	75.21	77.86	0.6562	0.6610	64.71	81.79	0.0887*	0.2029	81.05	73.01	0.2778	0.2314
Fun	87.39	75.56	0.1224	0.0300**	-	-	-	-	87.39	71.52	0.0389**	0.0120**

Therefore, while we may conclude that being amused correlates with significantly better results, not being amused does not accompany a worse performance than not being attentive or not finding the topic interesting, for instance. Moreover, we do not find any support for the claim that humor distracted the respondents, since the average levels of attention (3.12) and the number of students in the 'high attention' versus 'low attention' groups (18 and 8) are exactly the same for the control and the test group. Finally, we note that the worst performances belong to the eight students in the test group who gave low marks to the 'attention' variable (with an average quiz result of 62.50) and the five students in the control group who gave high marks to the 'humor' variable (with an average quiz result of 64.71). In conclusion, and consistent with IHPT, we add a qualification to our second hypothesis: instructional humor may enhance learning if learners are genuinely amused. Our results are suggestive, but as they

rely on the performance of a subset of only seven respondents, further research (at a larger scale) becomes imperative if our second hypothesis is to be confirmed.

Concluding remarks and further work

This paper has explored the potential of large language models, in particular OpenAI's GPT-4, to contribute to the field of economics education by incorporating computational humor into instructional texts. The results suggest that GPT-4 can be successful in producing adequate instructional humor in an entirely autonomous fashion and that respondents who find the instructional text amusing achieve significantly higher quiz results. However, our *prima-facie* findings are small-sample results, and merely open avenues for further testing and potential confirmation.

If confirmed, these results point toward at-scale, AI-driven personalization of instructional humor. The education sector – comprising both longer established institutions and edtechs – is actively looking for guidance on harnessing the power of AI to benefit each learner. Research going beyond the present study could be invaluable in this regard. Given a learner's characteristics, at what points is it best to inject AI-generated instructional humor? Are some humor types (analogy, hyperbole, irony, word play) better than others in particular contexts? In the future, we can expect virtual instructors delivering educational content with human-like realism. How does humor impact learning in such a setting? How will the effectiveness of AI-generated instructional humor change if it is founded on humor algorithms from an experienced comedy writer (Toplyn, 2023)? Further technological developments will make it possible to examine the impact of AI-generated humor on learning in far greater detail. For example, non-invasive and easy-to-use *electroencephalography* technology has already been used to obtain experimental subjects' focus dynamics at high temporal resolution, and can also be used to study enjoyment, anxiety, 'flow' state, memory formation etc. (Haruvi *et al.*, 2022).

It is worth addressing an important objection sometimes encountered to the very premise of instructional humor: learners should be motivated enough that 'sugar-coating' with humor is not necessary. We agree that there are some learners who are fully committed and do not require such help. Others, however, should not be neglected – especially as they are likely to include disproportionate numbers of the educationally underprivileged, uninspired by the notion of learning for its own sake. Further, even the most committed learner in a main field may be less committed when it comes to other subjects. Our paper argues that, for these learners, AI-generated humor may be helpful (or at least not harmful). Such 'sugar-coating' is a form of temptation-bundling, which has increasingly been shown to be beneficial in a variety of contexts; for example, Milkman *et al*. (2014) show gym attendance to increase if going to the gym is bundled with listening to engaging audiobooks. (But note that, while listening to audiobooks makes a workout more enjoyable without making it more impactful, instructional humor may not only make the learning process more enjoyable, but also provide additional insights into the material and/or make it more memorable.) Thus, judicious use of AI-generated instructional humor can point the way toward making previously forbidding subject matter accessible and enjoyable to many learners who would otherwise be left behind.

References

Ash, E. and Hansen, S. (2023) 'Text algorithms in economics', *Annual Review of Economics,* 15, pp.659–88.

Bolkan, S., Griffin, D. and Goodboy, A. (2018) 'Humor in the classroom: the effects of integrated humor on student learning', *Communication Education,* 67, pp.144–64.

Buttussi, F. and Chittaro, L. (2020) 'Humor and fear appeals in animated pedagogical agents: an evaluation in aviation safety education', *IEEE Transactions on Learning Technologies,* 13, 1, pp.63–76.

Celik, B. and Gundogdu, K. (2016) 'The effect of using humor and concept cartoons in high school ICT lessons on students' achievement, retention, attitude and anxiety', *Computers and Education*, 103, pp.144–57.

Crompton, H. and Burke, D. (2023) 'Artificial intelligence in higher education: the state of the field', *International Journal of Educational Technology in Higher Education,* 20, pp.1–22.

Erdogdu, F. and Cakiroglu, U. (2021) 'The educational power of humor on student engagement in online learning environments', *Research and Practice in Technology Enhanced Learning*, 16, 1, article 9.

Haruvi, A., Kopito, R., Brande-Eilat, N. … Furman, D. (2022) 'Measuring and modeling the effect of audio on human focus in everyday environments using brain-computer interface technology', *Frontiers in Computational Neuroscience*, 15, pp.1–17.

Milkman, K., Minson, J. and Volpp, K. (2014) 'Holding the Hunger Games hostage at the gym: an evaluation of temptation bundling', *Management Science,* 60, 2, pp.283–299.

Nair, V., Schumacher, E., Tso, G. and Kannan, A. (2023) 'DERA: enhancing Large Language Model completions with dialog-enabled resolving agents', available at https://www.researchgate.net/publication/369655617_DERA_Enhancing_Large_Language_Model_Completions_with_Dialog-Enabled_Resolving_Agents (accessed August 2024).

Ray, P. (2023) 'ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope', *Internet of Things and CyberPhysical Systems,* 3, pp.121–54.

Shinn, N., Cassano, F., Berman, E., Gopinath, A., Narasimhan K. and Yao, S. (2023) 'Reflexion: language agents with verbal reinforcement learning', available at https://openreview.net/forum?id=vAElhFcKW6 (accessed August 2024).

Toplyn, J. (2023) 'Witscript 3: a hybrid AI system for improvising jokes in a conversation', available at https://www.researchgate.net/publication/366983683_Witscript_3_A_Hybrid_AI_System_for_Improvising_Jokes_in_a_Conversation (accessed August 2024).

Wanzer, M., Frymier, A. and Irwin, J. (2010) 'An explanation of the relationship between instructor humor and student learning: instructional humor processing theory', *Communication Education,* 59, pp.1–18.

Warren, C., Barsky, A. and McGraw, A. (2020) 'What makes things funny? An integrative review of the antecedents of laughter and amusement', *Personality and Social Psychology Review,* 25, pp.41–65.

Appendices

Appendix A: Oliver Hart

Input text

Oliver Hart is a British-born economist who, along with Bengt Holmström, was awarded the Nobel Prize in Economics in 2016. Their main contributions are in the field of contract theory, which is basically about understanding how contracts and agreements work between people or parties in different situations. Hart specifically focused on the area called 'incomplete contracts'. In the real world, contracts between people or companies often cannot include every possible detail of a situation. There may be things left out or vague because people cannot predict everything that might happen in the future. Hart's work examines what happens when contracts are incomplete and how to deal with such situations. One of his main contributions is studying the role of ownership in contracts. He found that, when contracts are incomplete, ownership becomes a crucial factor in how decisions are made. For example, if a company decides to hire another organization to provide a service, it's important to determine who has control over the decision-making. Hart's work showed that having the right ownership structure can help solve issues that may arise from incomplete contracts. Hart also studied how ownership affects the performance and productivity of organizations. For instance, he found that, when workers are given some degree of ownership, such as shares in the company or the ability to influence decisions, it can motivate them to work harder and improve the overall performance of the organization. Overall, Oliver Hart's contributions helped us to better understand contracts, ownership, and the importance of decision-making in diverse situations. His research has influenced many areas of economics, such as corporate finance, organizational design, and public policy, and earned him the Nobel Prize. *(276 words)*

Output text: computational humor

Ah, Oliver Hart – the British-born economist who, rumour has it, could make even the most tedious contract sound like a thrilling mystery novel. Along with Bengt Holmström, he was awarded the Nobel Prize in Economics in 2016 for their work in the field of contract theory. You know, that riveting topic that keeps us all on the edge of our seats at dinner parties. Hart specifically focused on the area called 'incomplete contracts'. It's like trying to predict the weather – you can't possibly include every detail in a contract because the future is just too uncertain. Hart's work examines what happens when contracts are incomplete and how to navigate these murky waters. One of his main contributions is studying the role of ownership in contracts. Picture this: you're at a party, and there's only one slice of pizza left. Who gets it? That's where ownership comes in. Hart found that when contracts are incomplete, ownership becomes a crucial factor in decision-making. For example, if a company decides to hire another organization to provide a service, it's important to determine who gets to call the shots. Hart also studied how ownership affects the performance and productivity of organizations. Turns out, giving workers a piece of the pie (not just the pizza kind) can motivate them to work harder and improve the overall performance of the organization. Who knew a few shares could make such a difference? In a nutshell, Oliver Hart's contributions helped us better understand contracts, ownership, and the importance of decision-making in various situations. His research has influenced many areas of economics, such as corporate finance, organizational design, and public policy. And, of course, it earned him that shiny Nobel Prize – a well-deserved reward for making contracts a little less yawn-inducing. *(295 words)*

Appendix B: Paul Milgrom

Input text

Paul R. Milgrom, along with Robert B. Wilson, received the Nobel Prize in Economic Sciences in 2020 for their contributions to auction theory and inventions of new auction formats. In simple terms, Milgrom's work focused on understanding how auctions work and creating better ways to conduct them. Auctions are events where people bid on items or services, such as art, rare collectibles, government contracts, or even radio frequencies. Milgrom's contributions can be broken down into two main parts:

1 Auction Theory: Milgrom helped develop better ways to understand how people should bid in auctions based on their own value estimation of the item or service being auctioned, the auction format, and how they expect others to bid. This is important because auctions are commonly used in everyday life and have a significant impact on businesses, governments, and individuals.
2 New Auction Formats: Alongside his work on auction theory, Milgrom, in collaboration with Wilson, invented new auction formats to better allocate resources or services. For example, they designed the 'Simultaneous Multiple Round Auction' (SMRA), which is now widely used by governments to allocate radio frequencies to telecommunication companies. The goal of these new formats is to ensure that the auctioned items are allocated to the participants who value them the most, leading to more efficient outcomes.

In summary, Paul Milgrom's Nobel Prize-winning contributions lie in his innovative work on understanding and improving the way auctions work. His efforts have led to the development of new auction formats that are more efficient and effective in allocating resources and services. *(258 words)*

Output text: computational humor

Ah, auctions – the thrill of the hunt, the anticipation of snagging a deal, and the nail-biting suspense of outbidding your rivals! Our friend Paul R. Milgrom, along with Robert B. Wilson, scored the ultimate prize, the Nobel Prize in Economic Sciences in 2020, for their fantastic work in auction theory and creating brand-new auction formats. Now let's break it down, like an auctioneer counting down the final seconds:

1 Auction Theory: Milgrom's auction theory is like a treasure map for bidders. It helps them figure out the best bidding strategies based on their own value estimation, the auction format, and their predictions of how others will bid (cue dramatic music). It's important because, let's face it, auctions are everywhere – from eBay to government contracts, they shape our lives and economy.
2 New Auction Formats: Milgrom and Wilson didn't stop at auction theory. They went the extra mile and designed new auction formats, like the 'Simultaneous Multiple Round Auction' (SMRA). Picture a high-stakes, multi-table poker game, but instead of cards, it's radio frequencies auctioned for telecommunication companies. These innovative formats aim to make sure the auctioned items end up with those who value them the most, leading to more efficient outcomes.

So, Paul Milgrom's Nobel Prize-worthy contributions have given us a better understanding of the fascinating world of auctions. His work has not only spiced up the auction scene but also led to more efficient and effective ways of allocating resources and services. Talk about hitting the jackpot! *(249 words)*

Appendix C: Bengt Holmström

Input text

Bengt Holmström is a Finnish economist who has made significant contributions to the field of contract theory, specifically in the areas of information asymmetry and incentive mechanisms. One of Holmström's key contributions is the development of the informativeness principle, which states that optimal contracts should base performance measures on all relevant and informative signals about the actions taken by an agent, such as a CEO, in order to get the best understanding of their performance. This principle highlights the importance of designing contracts that carefully balance incentives and risks, taking into account both the agent's performance and the outcome of their actions. In his analysis on how a CEO's contract should be formulated, Holmström proposed the 'multi-tasking model', which acknowledges the complexity of a CEO's role and the various tasks they need to perform. According to his model, the contract should provide incentives that enable the CEO to prioritize and allocate resources effectively among different tasks. The contract should also ensure that the CEO focuses on long-term strategic goals instead of short-term financial gains, thereby avoiding potential conflicts of interest or moral hazard situations. Another major contribution of Holmström's work is the development of the 'career concern' model. This model emphasizes the importance of a CEO's reputation in the labor market and how it affects their contractual incentives. The career concern model suggests that a CEO's desire to build and maintain a strong reputation can motivate them to make decisions that are beneficial for the company, even if those decisions do not result in immediate financial gains. One of Holmström's significant contributions that led to his Nobel Prize win is his work on the 'hidden action' or 'moral hazard in teams' problem. He analysed how to design optimal contracts when the actions of individual team members cannot be easily observed, and their efforts cannot be precisely determined. Holmström's work resulted in the development of a linear sharing rule that addresses the free-rider problem within teams, thereby providing a framework for designing contracts that enhance teamwork and collaboration. This linear sharing rule helped to improve understanding of how incentive structures can impact the overall productivity and efficiency of organizations. *(358 words)*

Output text: computational humor

Bengt Holmström, a Finnish economist with a knack for tackling complex concepts, is our main focus today. He's a bit of a rock star in the world of contract theory, specifically in the areas of information asymmetry and incentive mechanisms. You could say he's the Sherlock Holmes of economics, always in search of informative clues. One of Holmström's key contributions is the development of the informativeness principle. Think of it as a detective story where the optimal contracts act like the magnifying glass, revealing all relevant and informative signals about the actions taken by an agent, such as a CEO. This way, we can have the clearest picture of their performance. The principle highlights the importance of designing contracts that are as carefully balanced as a tightrope walker, taking into account both the agent's performance and the outcome of their actions. When it comes to our multitasking maestros, the CEOs, Holmström came up with the aptly named 'multi-tasking model'. This model acknowledges that a CEO's role is as complex as a Rubik's cube, with various tasks that need to be tackled effectively. The contract should provide incentives that enable the CEO to prioritize and allocate resources like a pro juggler, without dropping the ball on long-term strategic goals in favor of short-term financial gains. That way, we can sidestep potential conflicts of interest or moral hazard situations. As we move on to the 'career concern' model, we see the spotlight shining on a CEO's reputation in the labor market. This model suggests that a CEO's desire to build and maintain a strong reputation is like a marathon runner aiming for that personal best; it can motivate them to make decisions that benefit the company, even if those decisions don't result in a finish line filled with immediate financial gains. Holmström's

grand finale, which contributed to his Nobel Prize win, is his work on the 'hidden action' or 'moral hazard in teams' problem. Picture a group project where everyone claims to be working hard, but it's unclear who's actually pulling their weight. Holmström's work led to the development of a linear sharing rule that tackles the free-rider problem within teams, like a referee keeping everyone in check. This rule provides a framework for designing contracts that enhance teamwork and collaboration, ultimately improving the overall productivity and efficiency of organizations. *(366 words)*

Data availability

The data used in this paper can be accessed at https://data.mendeley.com/datasets/rh34mry2xv/1.