

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☒ ☐ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Data was collected using the publicly available version of StarCraft II (versions 4.8.2 to 4.10), developed by Blizzard Entertainment.

Data analysis We used the open source environment to interact with the game of StarCraft II, provided by Blizzard and DeepMind (<https://github.com/deepmind/pysc2>), using the game version 4.10. The networks used the TensorFlow 1.0 library with custom extensions. Analysis was performed with custom code written in Python 2.7. We additionally provide pseudocode for all algorithms described in the paper.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

We did provide both the raw data used in the paper from the online experiment, and all the evaluation games played in the StarCraft II standard Replay format. The dataset containing all the replays used for imitation learning are distributed by Blizzard using a specific API: <https://github.com/Blizzard/s2client-proto>

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	To study our agents performance, we played a total of 360 games online against the population of players that play StarCraft II in the European servers. The sample size was determined with consultation with Blizzard and professional players, who deemed that 60 games would be sufficient to estimate performance of a new professional level player reliably with low uncertainty (less than 50 MMR). We did play 90 games total, per race, plus 30, per race, for supervised agents. For the league analysis we used around 130,000,000 full games of agent vs agent, and for ablations, we used around 20,000,000 games.
Data exclusions	No data was excluded from the study.
Replication	Because of the nature of the game, we did perform three independent experiments, using three distinct races. From the total of 9 runs, we did not observe any significant deviation, and thus we reproduced the intended conditions of the experiment ourselves. Because we played anonymously, reproducing the same conditions in future studies should be relatively easy, assuming care is taken to remain anonymous.
Randomization	The players and order in which we played against them was determined by the matchmaking algorithm that Blizzard employs to create matches in their online service, which was designed many years before our study, and which the authors of this manuscript had no control over. Such system is solely based on the skill level of players, and is thus random and the authors of this manuscript were blind to group allocation.
Blinding	The authors were blind to group allocation. See "Randomization".

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging