Electronic Supplementary material
Table of Contents

Appendix 1: Reproductive success

i) Stages of reproductive failure
Figure S1: The change in hatching success with age and boldness
Figure S2: The change in fledging success with age.
ii) Within vs between individual level changes in reproductive success with age
Table S1: The number of years of data for each individual included in the main reproductive success analyses.

Appendix 2: Boldness

i) Exploring boldness
Figure S3: Histogram of boldness scores
ii) Plasticity in boldness
Figure S4: The change in raw boldness measured with age. This plot does not attempt to partition repeated measures from changes with age.
Figure S5: The change in boldness between sequential observations within a single year.
Figure S6: The change in boldness with a change in age.

Appendix 3: Models fitting changes with age

Table S2: Model structure and AIC values for models using different age structures for population reproductive success and foraging trip duration.

Appendix 4: Results from the main analyses

Tables S3: Age at first reproduction models
Tables S4: Reproductive success across lifetime models
Tables S5: Early adulthood reproduction models
Tables S6: Late adulthood reproduction models
Table S7: Foraging trip duration model
Table S8: Mass gain model

Appendix 5: Cross sectional vs longitudinal analyses in reproductive success and foraging behaviour

Table S9: Parameter estimates for reproductive success models, using the cross sectional data, where boldness was measured in the same year as reproductive success.
Figure S7: Reproductive success, age and boldness using only cross sectional data, where boldness was measured in the same year as reproductive success.
Table S10: Parameter estimates for foraging models, using the cross sectional data, where boldness was measured in the same year as foraging trip duration.
Figure S8: Foraging, age and boldness using only cross sectional data, where boldness was measured in the same year as foraging trip duration.

Appendix 1

i) Stages of reproductive failure

As it is widely known in wandering albatrosses that most reproductive failures occur during incubation, we examined the main results, subdividing the attempts into hatching success and, for chicks that hatched, fledging success. For simplicity the results we present here use a single parameter estimate of boldness per bird from the Bayesian mixed model. Results for reproductive success were mimicked in hatching success, with strong differences in males ($\chi^2_1 = 5.08$; $p = 0.024$; Figure S1a) but not females ($\chi^2_1 = 0.30$; $p = 0.58$; Figure S1b). Fledging success was high across all individuals and explained only by age^2 (Age^2: $\chi^2_1 = 18.59$; $p < 0.001$; Figure S2).

Figure S1: The change in hatching success with age and boldness. A) Bolder males senesce less quickly than shyer males, with solid lines showing the significant interaction between boldness and age. B) Females do not show significant differences in hatching success senescence with boldness. For illustrative purposes dashed lines show non-significant differences in hatching success with age between the personality types. While age and boldness are continuous measures in all analyses, for illustrative purposes only, they are grouped here. Age was grouped into 5 year bins from 5 - 45 years and the mean reproductive success (± SE) plotted against the mid-point of the bin. Any reproductive attempt after 42 years of age was collapsed into to 40-45 year bin, in keeping with the analyses. Females are plotted in shades of red and males in shades of blue. Boldness was grouped into two categories: Bold (Upper 1/2 of boldness scores; Dark blue/dark red); Shy (Lower 1/2 of boldness scores; Pale blue/Pink).

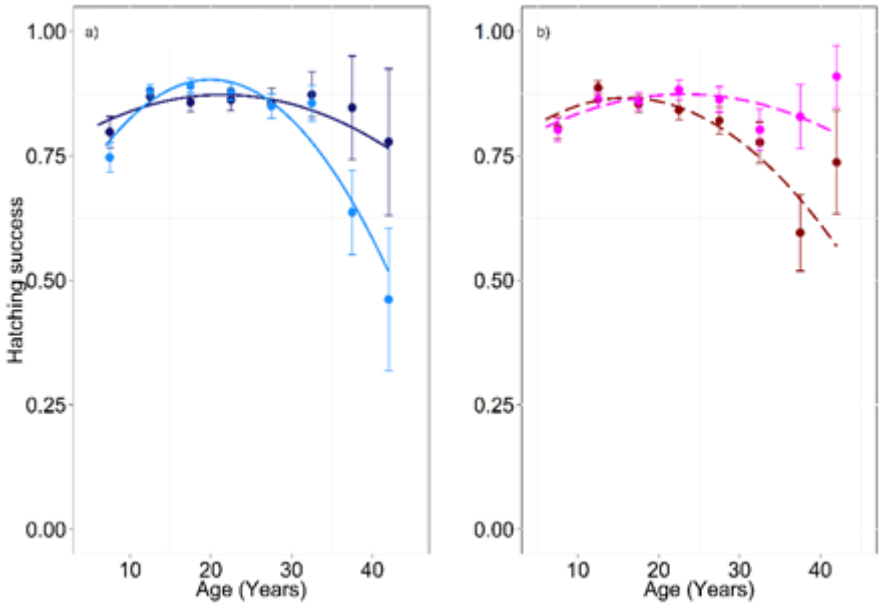
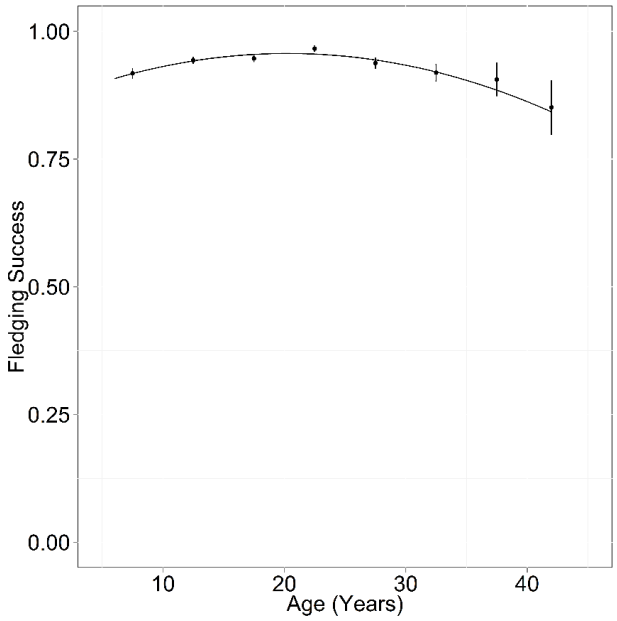


Figure S2: The change in fledging success with age. While age is a continuous measure in all analyses, for illustrative purposes only, it is grouped here. Age was grouped into 5 year bins from 5 - 45 years and the mean reproductive success (± SE) plotted against the mid-point of the bin. Any reproductive attempt after 42 years of age was collapsed into to 40-45 year bin, in keeping with the analyses.



ii) Within vs between individual level changes in reproductive success with age

In our data, while there are many birds with data across their entire lifetime, we also use data from some birds that have only been sampled across part of their lifetime (see Table S1 for exact numbers of individuals). This is inevitable in a species that lives for over 50 years. However, with this data structure comes some analytical issues. The main concern being that in many studies examining senescence, individual level variation in senescence rates has been measured. However, having sampled some individuals over only a subsection of ages, this method presents problems for our data. If we imagine an extreme scenario. Bird A: We have 5 reproductive attempts, measured between 5 and 15 years of age. Bird B: We have 5 reproductive attempts, measured between 15 and 25 years of age. A within individual level effect may show that bird A increases its reproductive performance with age; Bird B decreases its reproductive performance with age. If we designate this variation as between individual differences in senescence, we would be naïve. Had we the data for both birds between 5 and 25 years, we may in fact find no difference in their ageing.

A recent paper on senescence in wandering albatrosses found similar issues with the age ranges measured between individuals (Froy et al. 2013 Ecol. Letts). Here the authors addressed this problem by splitting the data into "early adulthood" and "late adulthood". This allowed within individual differences to be partitioned, including only birds with sufficient data. In the main results section we include an analysis where we divide the data into two groups: i) Reproductive attempts from birds <22 years old and with data available between breeding commencement and 20 years old or more ii) Reproductive attempts from birds ≥22 years old and with data available until at least 38 years of age. By restricting our

data set in this way, we remove any possibility that sampling regime may drive differences between individuals. This then allows us to fit an individual level senescent slope as a random slope in all these model. The strong congruence across models, and the low variance explained by individual level slopes, suggests within individual variation in senescence rates, above that which is explained by personality, is low.

Table S1: The number of years data for each individual included in the main reproductive success analyses. Some birds of known personality were excluded as there was insufficient data on their reproductive success (N = 175). This occurred as reproductive success was not known for 2013 at the time of writing but personality had been tested in this year.

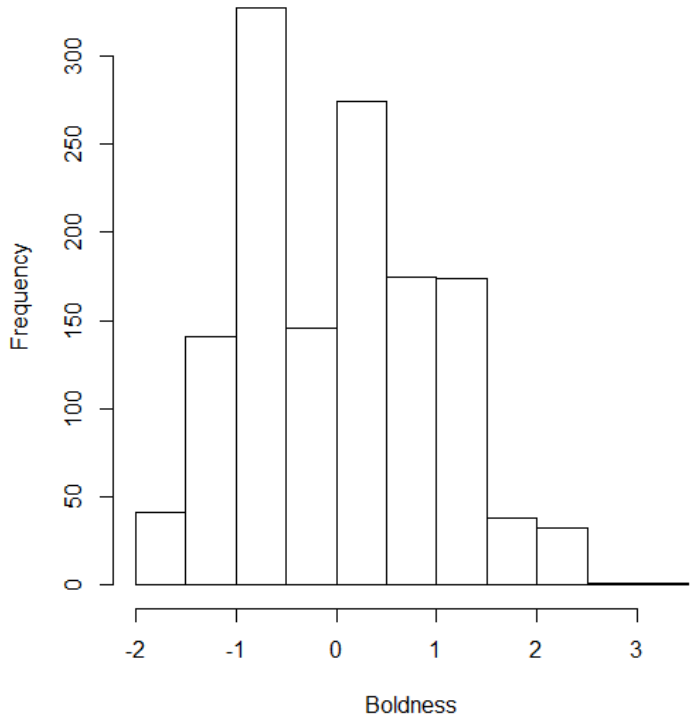
Number of Years data	Number of Individuals
1 - 5	288
5- 10	233
10- 15	254
15 - 20	171
20 - 25	126
25 - 30	43
30 - 35	14
35 - 40	7
40 - 50	6

Appendix 2

i) Exploring boldness

For all main analyses we used 1000 estimates of boldness per bird, to incorporate uncertainty into our models. Below is a plot of the mean boldness per bird, demonstrating the average distribution of scores. The maximal range across all estimates was -3.80 to 4.51.

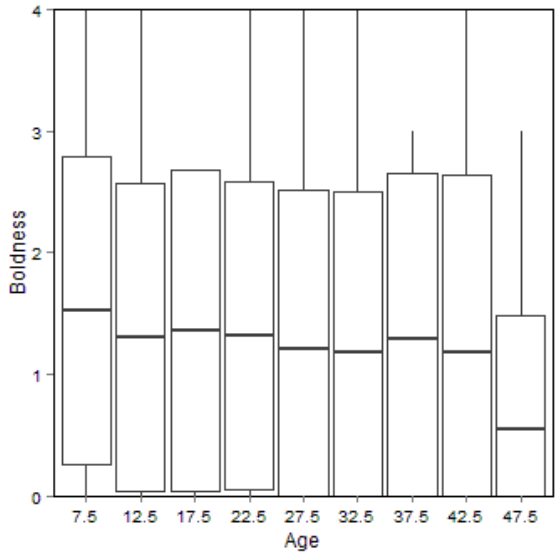
Figure S3: Histogram of boldness scores



ii) Plasticity in boldness

In a previous paper, we report that there is a difference in boldness across age groups, such that there a weak effect that birds are shyer when old. While this effect is very slight (See figure S4), here we address the question of whether this change is constant across individuals, or whether individuals exhibit within individual plasticity with age. When observing an individual multiple times, there are several factors to take into account. First, the number of times a bird has been tested always increases by one per measurement taken. These repeat observations may occur within a year, such that the age is constant or between years, and the age also increases with observation number.

Figure S4: The change in raw boldness measures with age. This plot does not attempt to partition repeated measures from changes with age. The box plot shows medians, quartiles and the range of values per age group.



Constraints on changes in boldness

First, we discuss briefly the constraints of using an ordinal score on the possible changes in behaviour with time. Any individual which had the maximum score (4) was constrained to remain the same or to become shyer with time. Conversely, any individual which had a score of 0, could only remain the same or become bolder with time. These changes can be seen in Figures S5 and S6 below. As such, 40% of the possible scores we have are by definition forced to have a slope of zero or slopes with opposite signs. This will appear in the data as birds who have an opposite response to the changing condition (observation number or age). Furthermore, individuals are forced to change boldness by units of 1. This makes us call into question any interpretation of "within individual plasticity" in this population from the data we have collected.

However, we conduct these analysis to allow us to examine whether individual changes occur more at the ends of the distribution.

Here we fit two models:

a) Allowing individuals to change their boldness between repeated observations within a single year
b) Allowing individuals to change their boldness between repeated observations between years, so that age also changes

We were unable to fit a model that attempts to allow individuals to vary within and between years differently, as we had insufficient data.

Using methods outlined in van der Pol and Verhulst (2006), we fitted the following basic models for each data set (MC = mean centred):

i) Population model : Rij = β0 + βs x Population MC Variable+ μ0i + e0ij

This examines the population level response to the change, in this case multiple observations within the same year or multiple observations at different ages.

ii) Individual model: Rij = β0 + βw x Individual MC Variable + βs x Population MC Variable + μ0i + e0ij

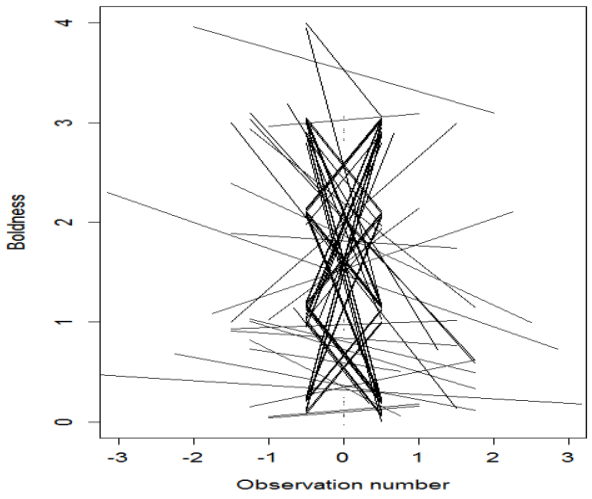
This extends the population model, to allow each individual to have its own response to the change.

a) Observation effects (within a year)

i) We began with a population level model, which fitted observation number as a fixed effect and individual ID as a random intercept. This showed a significant population level change in boldness. This demonstrates that birds on average become shyer with multiple tests within the same year (Estimate = -0.13 ± 0.00). We account for this in our main models.

ii) We then extended this analysis to allow individuals to change their boldness individually with observation number. We then repeated the model, including a random slope per individual, to allow us to categorise within individual plasticity effects (χ^2_2 = 280.36; p < 0.001). While the results suggest a strong within individual level change in boldness with repeated observations, you can see that because if birds change boldness they must change by 20% because the data is ordinal and, as explained previously, birds at the limits of the score are constrained in their change.

Figure S5: The change in boldness between sequential observations within a single year.



Scores are scaled for observer differences and most individuals change by one unit only; the minimum measureable change. This could be evidence of individual plasticity or it could be evidence of variation in measurement. We feel very strongly that to assess plasticity in a given direction, 3 or more measurements are necessary to ensure variation in sampling does not get erroneously partitioned as plasticity.

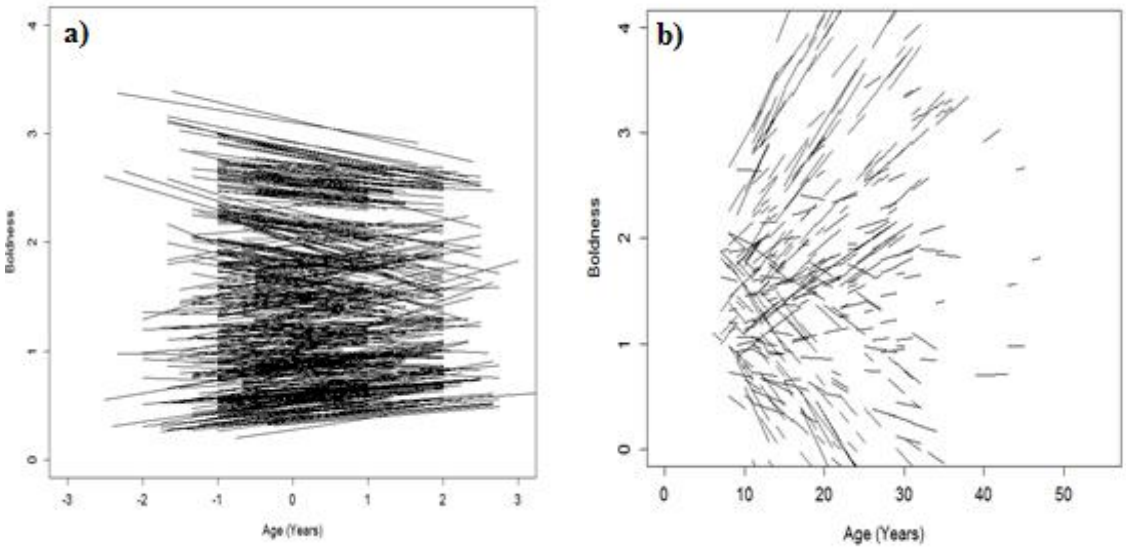
As we have measurements between years per bird, this allows us to try to split changes with age from changes with subsequent observations.

b) Age effects (between years)

i) We began with a population level model, which fitted observation number as a fixed effect and individual ID as a random intercept. Using only repeats, there was no significant change in boldness with age ($\chi^2_1 = 2.21$; $p = 0.14$). Data previously reported in Patrick et al. 2013 included all individuals.

ii) We then extended this analysis to allow individuals to change their boldness individually with age. We then repeated the model, including a random slope per individual, to allow us to categorise within individual plasticity effects ($\chi^2_2 = 9.21$; $p = 0.01$). These individual effects were much smaller than those seen between observations but within years and if you explore these slopes more thoroughly you can see they are very shallow.

Figure S6: The change in boldness with a change in age. The graph show little evidence that individual differences with age are greater than those between sequential tests, aside from those where birds had an original score of zero or four. Furthermore there is no evidence of changes in plasticity with age at the population level. Figure S6a: Individual slopes plotted against mean centred age. Figure S6b: Individual slopes plotted against actual age.



In summary, when you partition between year from between observation changes, the changes in boldness with age is negligible in comparison to changes within a year. While these models have shown some evidence of changes within birds between observations, we feel this is unlikely to be plasticity between birds as these differences do not seem to persist in between years, where we have more information on each individual (single year scores are based on more than one observation). Repeated observations were collected to allow us to estimate boldness with improved precision. Our score has a repeatability of nearly 0.5 which suggest that 50% can be explained by individuals consistently behaving in the same way. For this paper, we feel it is appropriate to focus on the consistency of this measure, and until further data is available, we feel unable to categorise individual plasticity in boldness from measurement differences in the wild.

Appendix 3: Comparing the structure of age related changes in reproductive success and foraging behaviour

Table S2: The population reproductive success and foraging trip duration modelled using different age related changes. Quadratic models had the lowest AIC for both males and females for reproductive success. For foraging behaviour, although there was slightly stronger support for a quadratic age effect, these models converged poorly, suggesting overfitting, so a linear age effect was investigated across the birds lifetime, and in early and late adulthood (which mimics a quadratic effect).

Model formula	**Biological meaning**	**Males**	**Females**
RS ~ 1	No effect of age on reproductive success	3801.519	4100.004
RS ~ Age	Linear effect of age on reproductive success	3802.978	4101.541
RS ~ Age^2 + Age	**Quadratic effect of age on reproductive success**	**3749.181**	**4067.882**
RS ~ T1.Age + T2.Age	Existence of a threshold effect of age on reproductive success (1 threshold)	3796.079	4083.847
RS ~ T1.Age + T2.Age + T3.Age	Existence of a threshold effect of age on reproductive success (2 threshold)	3780.961	4080.437
Duration of foraging trip ~ 1	No effect of age on the duration of foraging trip	895.8338	1016.868
Duration of foraging trip ~ Age	**Linear effect of age on the duration of foraging trip**	**891.4357**	**1012.67**
Duration of foraging trip ~ Age^2 + Age	Quadratic effect of age on the duration of foraging trip	885.1067	1005.066
Duration of foraging trip ~ T1.Age + T2.Age	Existence of a threshold effect of age on the duration of foraging trip (1 threshold)	899.9986	1019.873
Duration of foraging trip~ T1.Age + T2.Age + T3.Age	Existence of a threshold effect of age on the duration of foraging trip (2 threshold)	904.1002	1024.726

Appendix 4: Results tables from the main models

Age at first reproduction models

Table S3: Parameter estimates for age at first reproduction models. Estimates are on a log link scale and bold font is used for significant parameters in final models and non-significant interactions, which were dropped from the model, are underlined. Main model = Boldness * Sex. The range of estimates are the 95% confidence intervals extracted from a model that uses 1000 estimates of boldness per individual, examining the effect of this uncertainty in boldness on final parameter estimates.

Parameter Estimate			
Fixed Effects	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**
Intercept (Sex F)	**2.17±0.01**	**χ^2_1 = 26.09; p<0.001**	**2.17 - 2.17**
Sex M	0.10±0.02		0.10 - 0.10
Boldness	0.00±0.01	χ^2_1 = 0.30; p = 0.59	0.00 - 0.01
<u>Boldness x Sex</u>	<u>-0.01±0.02</u>	<u>χ^2_1 = 0.32; p = 0.57</u>	<u>-0.01 - 0.01</u>

Reproductive success across lifetime models

Table S4: Parameter estimates for full reproductive success models, using the full data set. Estimates are on a logit link scale and bold font is used for significant parameters and non-significant interactions, which were dropped from the model, are underlined. Main model = Boldness * Age^2 + Colony + Age at first reproduction. Colony abbreviations: BA = Baie Américane; BDN = Baie du Marin Nord; BDS = Baie du Marin Sud; CN = Chaloupe Nord; CS = Chaloupe Sud; Isle = Isle; PB = Pointe Basse; SN = Sphinx Nord; SS = Sphinx Sud. The range of estimates are the 95% confidence intervals extracted from a model that uses 1000 estimates of boldness per individual, examining the effect of this uncertainty in boldness on final parameter estimates.

Parameter Estimate	**Males**			**Females**		
Fixed Effects	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**
Intercept (BA Colony)	1.84±0.32	$\chi^2_9 = 16.25$ p = 0.06	1.80 - 1.86	0.97±0.34	$\chi^2_9 = 15.14$; p = 0.09	0.92-0.98
BDN Colony	0.19±0.26		0.18 - 0.22	0.74±0.29		0.74-0.77
BDS Colony	-0.11±0.22		-0.13 - -0.09	0.04±0.25		0.04-0.06
CN Colony	-0.10±0.37		-0.13 - -0.05	0.13±0.38		0.12-0.16
CS Colony	0.28±0.29		0.25 - 0.29	0.35±0.31		0.34-0.36
Isle Colony	-1.25±0.48		-1.27 - -1.21	-0.03±0.47		-0.04-0.00
PB Colony	-0.06±0.21		-0.07 - -0.04	0.05±0.23		0.04-0.08
PJ Colony	0.49±0.35		0.47 - 0.53	0.12±0.34		0.12-0.15
SN Colony	-0.01±0.37		-0.02 - 0.03	-0.11±0.34		-0.11--0.08
SS Colony	0.27±0.28		0.25 - 0.29	0.31±0.3		0.31-0.34
Boldness	-0.04±0.04		-0.06 - 0.02	-0.06±0.04	$\chi^2_1 = 1.70$; p = 0.19	-0.07-0.02
Age	1.71±4.48		1.57 - 3.9	-2.18±3.9		-2.27--1.93
Age at first reproduction	**-0.06±0.02**	**$\chi^2_1 = 5.21$ p = 0.02**	**-0.06 - -0.05**	0.04±0.03	$\chi^2_1 = 1.91$; p = 0.17	0.04-0.04
Age^2	-30.36±3.95		-31.11 - -28.22	**-19.94±3.32**	**$\chi^2_1 = 35.18$; p<0.001**	**-20.41--19.94**
Boldness x Age	2.37±3.24		-0.91 - 6.16	-4.55±2.93	$\chi^2_1 = 2.39$; p = 0.12	-5.34 - 2.81
Boldness x Age^2	**8.37±2.96**	**$\chi^2_1 = 7.98$ p = 0.005**	**3.04 - 9.51**	-1.06±2.71	$\chi^2_1 = 0.15$; p = 0.70	-3.93 - 2.7
Random Effects						
Bird ID	0.31±0.56		0.30 - 0.33	0.46±0.68		0.46 - 0.47
Year	0.10±0.32		0.10 - 0.11	0.02±0.15		0.02 - 0.02

Early adulthood reproductive success models

Table S5: Parameter estimates for reproductive success models, using only data before the age of onset for senescence (< 22 years old). Estimates are on a logit link scale and bold font is used for significant parameters and non-significant interactions, which were dropped from the model, are underlined. Main model = Boldness * Age + Colony + Age at first reproduction. Colony abbreviations: BDM = Baie Américane + Baie du Marin Nord + Baie du Marin Sud + Chaloupe Nord + Chaloupe Sud + Isle + Sphinx Nord + Sphinx Sud; PB = Pointe Basse. The range of estimates are the 95% confidence intervals extracted from a model that uses 1000 estimates of boldness per individual, examining the effect of this uncertainty in boldness on final parameter estimates.

	Males			**Females**		
Parameter Estimate	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**
Fixed Effects						
Intercept (BDM Colony)	1.56±0.39	$\chi^2_1 = 0.18$; p = 0.68	1.51-1.58	0.86±0.43	$\chi^2_1 = 1.28$; p = 0.26	0.79-0.88
PB Colony	0.06±0.14		0.04-0.10	-0.16±0.15		-0.18--0.13
Boldness	-0.10±0.06	$\chi^2_1 = 2.27$; p = 0.13	-0.12-0.00	-0.05±0.06	$\chi^2_1 = 0.86$; p = 0.35	-0.09-0.03
Age	**0.08±0.02**	**$\chi^2_1 = 16.80$; p < 0.001**	**0.08-0.08**	**0.04±0.02**	**$\chi^2_1 = 5.42$; p = 0.02**	**0.04-0.04**
Age at first reproduction	-0.03±0.04	$\chi^2_1 = 0.46$; p = 0.50	-0.03--0.02	0.01±0.04	$\chi^2_1 = 0.08$; p = 0.77	0.01-0.02
Boldness * Age	0.00±0.01	$\chi^2_1 = 0.10$; p = 0.76	-0.01-0.01	-0.01±0.01	$\chi^2_1 = 1.25$; p = 0.26	-0.01-0.00
Random Effects						
Bird ID Intercept	0.26±0.51		0.25-0.28	0.37±0.60		0.36-0.37
Bird ID Slope	0.00±0.07	$\chi^2_1 = 0.17$; p = 0.68	0.00-0.00	0.00±0.06	$\chi^2_1 = 3.16$; p = 0.21	0.00-0.00
Year	0.12±0.34		0.11-0.12	0.00±0.01		0.00-0.01

Late adulthood reproduction models

Table S6: Parameter estimates for reproductive success models, using only data collected after the age of onset for senescence (≥ 22 years old). Estimates are on a logit link scale and bold font is used for significant parameters and non-significant interactions, which were dropped from the model, are underlined. Main model = Boldness * Age + Colony + Age at first reproduction. Colony abbreviations: BDM = Baie Américane + Baie du Marin Nord + Baie du Marin Sud + Chaloupe Nord + Chaloupe Sud + Isle + Sphinx Nord + Sphinx Sud; PB = Pointe Basse. The range of estimates are the 95% confidence intervals extracted from a model that uses 1000 estimates of boldness per individual, examining the effect of this uncertainty in boldness on final parameter estimates.

Parameter Estimate	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**
Fixed Effects						
Intercept (BDM Colony)	-0.74±0.99	$\chi^2_1 = 0.78$; p = 0.38	-0.94 - -0.38	2.20±1.36	$\chi^2_1 = 1.83$; p = 0.18	1.36 - 2.16
PB Colony	0.54±0.61		0.28 - 0.69	-0.84±0.61		-1.10 - -0.65
Boldness	-0.06±0.16		-0.22 - 0.14	**-0.52±0.23**	**$\chi^2_1 = 5.04$; p = 0.02**	**-0.64 - -0.06**
Age	-0.11±0.04		-0.13 - -0.08	**-0.06±0.03**	**$\chi^2_1 = 5.48$; p = 0.02**	**-0.06 - -0.06**
Age at first reproduction	0.10±0.07	$\chi^2_1 = 2.58$; p = 0.11	0.08 - 0.12	0.16±0.09	$\chi^2_1 = 3.43$; p = 0.06	0.16 - 0.23
Boldness * Age	**0.05±0.02**	**$\chi^2_1 = 5.49$; p = 0.02**	**0.02 - 0.08**	0.01±0.03	$\chi^2_1 = 0.02$; p = 0.87	-0.01 - 0.04
Random Effects						
Bird ID Intercept	0.00±0.00		0.00 - 0.01	1.23±1.11		1.09 - 1.69
Bird ID Slope	0.00±0.08	$\chi^2_1 = 0.95$; p = 0.62	0.00 - 0.12	0.00±0.05	$\chi^2_1 = 1.80$; p = 0.41	0.00 - 0.07
Year	0.66±0.81		0.66 - 0.86	0.03±0.18		0.02 - 0.04

Foraging trip duration model

Table S7: Parameter estimates from models examining the duration of foraging trips across all ages. Bold font is used for significant parameters and non-significant interactions, which were dropped from the model, are underlined. Main model = Boldness * Age + Sex. The range of estimates are the 95% confidence intervals extracted from a model that uses 1000 estimates of boldness per individual, examining the effect of this uncertainty in boldness on final parameter estimates.

	All life			**Early Life**			**Late Life**		
Parameter Estimate	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**
Fixed Effects									
Intercept (Sex F)	9.93±1.06	$\chi^2_1 = 2.18$; p = 0.14	9.91 - 10.05	8.08±1.35	$\chi^2_1 = 0.59$;p = 0.44	8.00 - 8.14	10.32±1.36		9.79 - 10.37
Sex (M)	-0.89±0.6		-0.96 - -0.75	-0.57±0.74		-0.65 - -0.47	- 1.42±0.98	$\chi^2_1 = 1.91$; p = 0.17	- 1.53 - -1.08
Boldness	-0.03±0.27		-0.22 - 0.23	-0.15±0.36	$\chi^2_1 = 0.18$; p = 0.67	- 0.39 - 0.20	- 0.99±0.72		- 1.23 - -0.11
Age	0.27±5.56		-1.02 - 1.66	**-43.12±16.8**	**$\chi^2_1 = 6.65$; p = 0.01**	**-44.71 - -41.69**	7.22±13.66		7.20 - 14.65
Boldness x Age	**9.49±4.48**	**$\chi^2_1 = 4.52$;p = 0.03**	**1.47 - 10.05**	-1.27±0.72	$\chi^2_1 = 0.00$; p = 0.97	-13.58 - 10.72	**21.79±10.69**	**$\chi^2_1 = 4.29$; p = 0.04**	**5.83 - 24.62**
Random Effects									
Year	2.01±1.42		1.93 - 2.13	2.51±1.58		2.37 - 2.66	2.61±1.62		2.37 - 2.02
Month	2.30±1.52		2.07 - 2.62	2.55±1.60		2.37 - 2.76	1.22±1.10		0.61 - 2.02

Mass gain model

Table S8: Estimates of mass gain for individuals during foraging trips. Bold font is used for significant parameters and non-significant interactions, which were dropped from the model, are underlined. Main model = Boldness * Age + Sex. The range of estimates are the 95% confidence intervals extracted from a model that uses 1000 estimates of boldness per individual, examining the effect of this uncertainty in boldness on final parameter estimates.

	All life			**Early Life**			**Late Life**		
Parameter Estimate	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**	**Estimate ± SE**	**Test statistic and p value**	**Range of Estimates**
Fixed Effects									
Intercept (Sex F)	707.13±372.81	$\chi^2_1 = 1.43$; p = 0.23	628.1-728.24	1433.27±516.42	$\chi^2_1 = 0.48$; p = 0.49	1223.36-1495.41	400.08±740.96	$\chi^2_1 = 1.92$; p = 0.17	146.96-623.8
Sex (M)	-209.15±167.31		-240.19--180.41	165.31±245.67		118.97-186.32	-361.06±242.01		-428.59--266.56
Boldness	143.95±75.97	$\chi^2_1 = 3.68$; p = 0.06	32.77-158.96	90.2±111.33	$\chi^2_1 = 0.82$; p = 0.37	-87.94-123.97	**256.17±118.44**	**$\chi^2_1 = 5.34$; p = 0.02**	**91.57-293.21**
Age	7.51±8.09	$\chi^2_1 = 0.74$; p = 0.39	6.84-10.16	-37.73±34.36	$\chi^2_1 = 1.26$; p = 0.26	-43.37--22.14	20.84±20.07	$\chi^2_1 = 0.96$; p = 0.33	14.34-27.89
Boldness x Age	5.79±7.73	$\chi^2_1 = 0.64$; p = 0.42	-1.52-11.23	-22.40±26.06	$\chi^2_1 = 0.83$; p = 0.36	-61.49--5.51	0.52±18.90	$\chi^2_1 = 0.01$; p = 0.92	-10.49-18.00
Random Effects									
Year	0.00±0.00		0.00 - 138739208	0.00±0.00		0.00-11779	0.00±0.00		0.00-0.00
Month	237175±487		10011291 - 1522412198	11220±105.90		3164-39018.1009	148839±385.8		74764-261101

Appendix 5: Cross sectional vs longitudinal analyses

While longitudinal analyses are incredibly useful, as they allow the changes in behaviour over time to be measured, as we have only reproductive success measures longitudinally and boldness measured relatively cross-sectionally, here we explore the relationships using only cross sectional data. We fit identical models, but rather than using all reproductive estimates per bird, we use only those collected in years when birds had been sampled for boldness. We use identical boldness scores, controlling for observation number and observer. We use these in preference to raw scores in a given year. This is because we have shown here that there are clear differences among observers and that collecting several measures per bird allows us to better estimate actual personality. Therefore reproductive attempts between the year of first and last testing were used in these models. For example, if bird A was tested in year 1 and 3, a cumulative score of boldness is used from both measurements at each reproductive attempt. Foraging data is also relatively cross sectional but to be conservative, we repeat those analyses using truly cross-sectional data here.

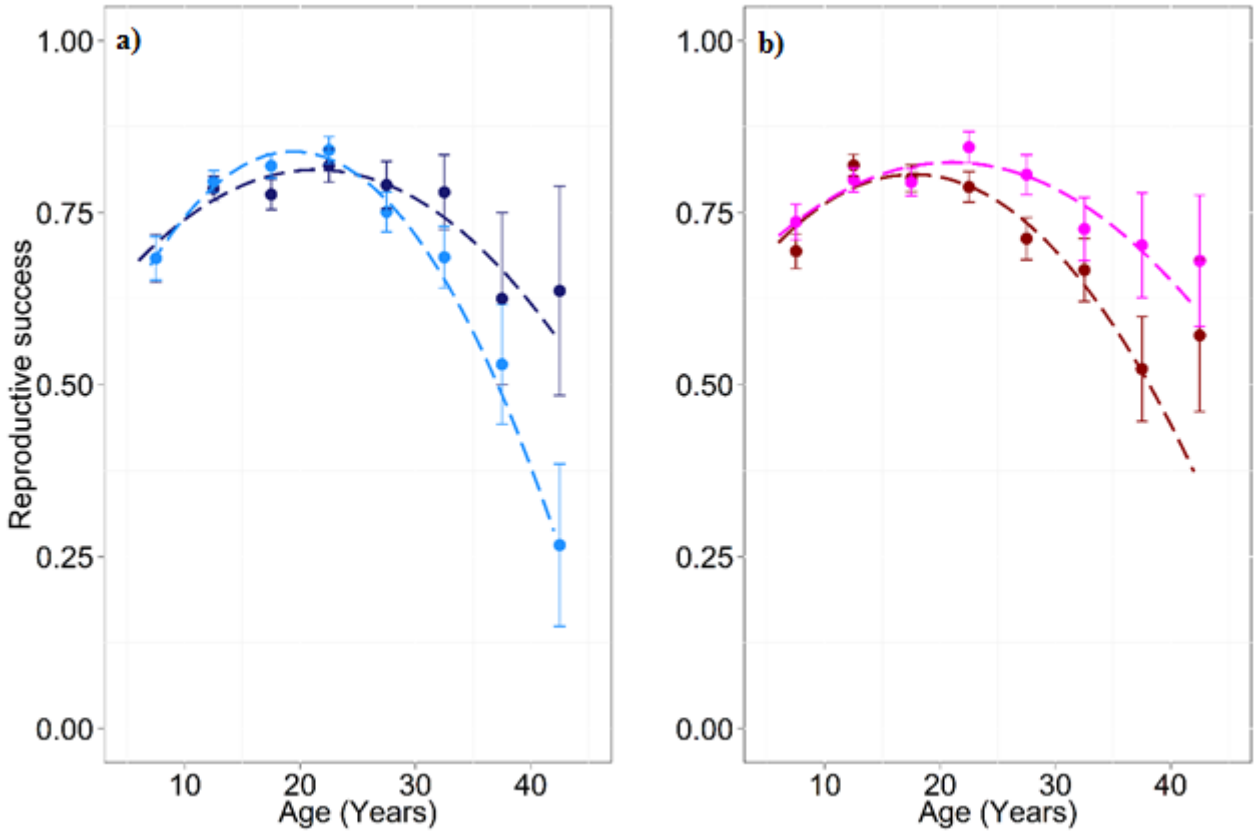
Reproductive success: Cross sectional analysis

While there appears insufficient power to detect a significant quadratic effect, effects of boldness on reproductive success with age are in the same direction as the main analysis (Table S9) and plots show very similar patterns (Figure S7).

Table S9: Parameter estimates for reproductive success models, using the cross sectional data only. Estimates are on a logit link scale and bold font is used for significant parameters and non-significant interactions, which were dropped from the model, are underlined. Colony abbreviations: BDM = Baie Américane + Baie du Marin Nord + Baie du Marin Sud + Chaloupe Nord + Chaloupe Sud + Isle + Sphinx Nord + Sphinx Sud; PB = Pointe Basse.

Parameter Estimate	**Males**		**Females**	
Fixed Effects	**Estimate ± SE**	**Test statistic and p value**	**Estimate ± SE**	**Test statistic and p value**
Intercept (BDM Colony)	2.09±0.59	$\chi^2_1 = 0.75$; p = 0.39	1.19±0.52	$\chi^2_1 = 0.13$; p = 0.72
PB Colony	-0.2±0.24		0.08±0.22	
Boldness	-0.11±0.1		-0.14±0.09	$\chi^2_1 = 2.33$; p = 0.13
Age	2.45±3.98		3.29±3.37	
Age at first reproduction	-0.05±0.06	$\chi^2_1 = 0.81$; p = 0.37	0.05±0.06	$\chi^2_1 = 0.69$; p = 0.40
Age^2	**-13.58±3.94**	**$\chi^2_1 = 11.66$; p<0.001**	**-13.88±3.09**	**$\chi^2_1 = 19.47$; p<0.001**
Boldness x Age	5.95±3.05	**$\chi^2_1 = 3.89$; p = 0.05**	2.60±2.60	$\chi^2_1 = 1.01$; p = 0.31
Boldness x Age^2	2.40±3.10	$\chi^2_1 = 0.62$; p = 0.43	-1.74±2.52	$\chi^2_1 = 0.48$; p = 0.49
Random Effects				
Year	0.00±0.04		0.00±0.00	

Figure S7: Reproductive success, age and boldness using only cross sectional data, where boldness was measured in the same year as reproductive success. A) Males B) Females. Results support those in the main paper. For illustrative purposes dashed lines show non-significant differences in reproductive success with age between the personality types. While age and boldness are continuous measures in all analyses, for illustrative purposes only, they are grouped here. Age was grouped into 5 year bins from 5 - 45 years and the mean reproductive success (± SE) plotted against the mid- point of the bin. Any reproductive attempt after 42 years of age was collapsed into to 40-45 year bin, in keeping with the analyses. Females are plotted in shades of red and males in shades of blue. Boldness was grouped into two categories: Bold (Upper 1/2 of boldness scores; Dark blue/dark red); Shy (Lower 1/2 of boldness scores; Pale blue/Pink).



Foraging trip duration: Cross sectional analysis

Table S10: Parameter estimates for foraging models, using the cross sectional data, where boldness was measured in the same year as reproductive success. Bold font is used for significant parameters and non-significant interactions, which were dropped from the model, are underlined.

	All life		**Early Life**		**Late Life**	
Parameter Estimate	**Estimate ± SE**	**Test statistic and p value**	**Estimate ± SE**	**Test statistic and p value**	**Estimate ± SE**	**Test statistic and p value**
Fixed Effects						
Intercept (Sex F)	**11.17±1.84**	**χ^2_1 = 4.42; p = 0.04**	7.82±2.77	χ^2_1 = 2.28; p = 0.13	13.03±2	χ^2_1 = 3.98; p = 0.05
Sex (M)	-1.87±0.89		-1.74±1.14		-2.71±1.39	
Boldness	-0.03±0.42		-0.8±0.53	χ^2_1 = 2.35; p = 0.13	-1.43±1.27	
Age	-1.08±5.73		**-39.59±19.1**	**χ^2_1 = 4.36; p = 0.04**	-5.7±15.01	
Boldness x Age	**14.66±5.19**	**χ^2_1 = 8.06; p = 0.005**	3.41±15.63	χ^2_1 = 0.05; p = 0.83	**32.49±14.8**	**χ^2_1 = 5.06; p = 0.02**
Random Effects						
Year	6.46±2.54		9.60±3.10		4.84±2.20	
Month	2.54±1.60		7.00±2.65		0.00±0.00	

Figure S8: Foraging, age and boldness using only cross sectional data, where boldness was measured in the same year as foraging data. Results support those in the main paper. While age and boldness are continuous measures in all analyses, for illustrative purposes only, they are grouped here. Age was grouped into 5 year bins from 5 - 45 years and the mean foraging duration (± SE) plotted against the mid- point of the bin. Any foraging trip after 42 years of age was collapsed into to 40-45 year bin, in keeping with the analyses. Boldness was grouped into two categories: Bold (Upper 1/2 of boldness scores; Dark green); Shy (Lower 1/2 of boldness scores; Green).